THE CONCEPTION OF SURPLUS IN THEORETICAL ECONOMICS

THE CONCEPTION OF SURPLUS IN THEORETICAL ECONOMICS

BY

AMIYAKUMAR DASGUPTA, M.A., PH.D.

LECTURER IN ECONOMICS IN THE UNIVERSITY OF DACCA

DASGUPTA & CO.

BOOK-SELLERS AND PUBLISHERS

CALCUTTA

1942

PRINTED BY
KALIPADA MUKHERJEE
AT EKA PRESS
210, CORNWALLIS STREET, CALCUTTA.

PUBLISHED BY
DASGUPTA & CO.
BOOK-SELLERS AND PUBLISHERS
54-3, COLLEGE STREET, CALCUTTA.

TO

THE MEMORY

OF

MY PARENTS

PREFACE

THE following study is the result of two years' research work in the London School of Economics and Political Science. The notion of Surplus with which it deals forms an extremely significant part of economic discussions. In a sense it is all-pervading. Since the beginning of economic science different writers have considered the problem of surplus from different points of view, and have sought for it in different contexts. I have tried to put them in their proper places and to clarify thought on these matters.

The study aims at an analysis of a concept, and as such it has been critical and interpretative. Problems have been sorted out from the general body of economic analysis for interpretation and criticism. And in the process quite an extensive field of economic enquiry has been covered. It may be hoped that the discussions that follow will be of some use towards an appreciation of certain implications of modern economic theory.

I should mention that it has not been possible for me to consider in the chapters on Consumer's Surplus the recent tendency towards a rehabilitation of the theory, as sponsored by Professors Hotelling and Hicks. This has been an act of omission, and I regret it. But I do not think that these developments would necessitate my withholding any of the observations made in the text.

Publication of the work has been delayed beyond anticipation. It was submitted and approved for Ph.D. degree of the University of London in the Michaelmas Term of the session 1936-37. Certain chapters—particularly the last two, dealing with Dynamic Surplus—have since been retouched, and two appendices added. These were published as articles in the *Indian Journal of Economics* (On

the Ultimate Nature of Costs, *I. J. E.*, Oct. 1939; The Elasticity of Reciprocal Demand and Terms of International Trade, *I. J. E.*, Oct. 1941), and have been reprinted with the kind permission of the Editors of the Journal.

A few words need specially to be said concerning the solution given in the Appendix on *Elasticity of Reciprocal Demand and Terms of International Trade.* It arose out of my reflections on a paper read before the Dacca University Economics Association, and purports to reconcile two opposing views on the question,—one given by Marshall and the other by Graham and Viner. Subsequent to the publication of my conclusions in the Journal, however, my attention was drawn to two articles (A. J. Duncan, Marshall's Paradox and the Direction of Shift of Demand, *Econometrica*, Vol. 6; and D. H. Robertson, Changes in International Demand and Terms of Trade, *Quarterly Journal of Economics*, Vol. 52) in which I found myself substantially anticipated. The solution nevertheless finds a place in the present volume, because the details with which it has been set out might, it is hoped, interest the general reader.

I should express my deepest gratitude to Professor Lionel Robbins for the constant care and attention that he showed while the work was in progress. I am also greatly indebted to Dr. P. N. Rosenstein-Rodan who read the entire manuscript and gave me valuable suggestions. In general I learnt much from essays and discussions in the Economic Theory Seminar of Professors Lionel Robbins and F. von Hayek, and from the lectures on Advanced Economic Theory held in the London School of Economics and Political Science during the sessions 1934-36.

RAJSA
July, 1942

A. K. DASGUPTA

CONTENTS

CHAPTER		PAGE
I.	Introduction and Plan	1
	DISTRIBUTIONAL SURPLUS STATIC ASPECT	
II.	The Theory of Surplus in Relation to Prosperity ..	17
III.	A Critical Examination of the Theory of Surplus Value and Unearned Income	30
IV.	Labour Cost and the Differential Surplus Theory of Rent	48
V.	Cost of Production and Rent Further Considered	68
Appendix.	On the Ultimate Nature of Costs ..	88
	GAINS FROM EXCHANGE CONSUMER'S SURPLUS	
VI.	Gains from Exchange: Earlier Notions ..	101
VII.	Marshall on Consumer's Surplus	117
VIII.	"Consumer's Surplus" in International Trade ..	138
Appendix.	On the Elasticity of Reciprocal Demand and Terms of International Trade	156
	DISTRIBUTIONAL SURPLUS DYNAMIC ASPECT	
IX.	Frictional Surplus	171
X.	Money and Dynamic Surplus	189
Index of Authors Cited		213

CHAPTER I

INTRODUCTION AND PLAN

1. In the present study an attempt will be made to examine the conception of Surplus in Economic Theory. We shall see the nature of different categories of Surplus with reference to their proper contexts, and examine what degree of significance can be attached to them.

The concept of surplus presupposes the existence of two entities which are taken for comparison. Whenever we speak of surplus we refer to a standard with which a particular entity is compared. When, for example, we speak of 'surplus stocks' we mean an amount which is in excess of that which has an effective demand. While speaking of 'surplus population' we mean a number of people that cannot be profitably employed and cannot have proper subsistence; or, perhaps, we take other standards of comparison with reference to which some people are found redundant. Similarly we may speak of a 'surplus budget' where the public revenue exceeds public expenditure for a financial year; and so on.

In the present context the conception of surplus will be used in the sense of a comparison between Cost and Return. Surplus is an excess of 'return' over 'costs'. But what is 'return' and what are costs, and from whose point of view are they to be considered? Costs and surplus may relate to the producer, or they may relate to the consumer. From the point of view of the producer, that is to say, from the point of view of the resources employed in production, surplus is any amount by which the income actually

secured may be exceeding the income that is sacrificed. From the point of view of the consumer spending money his surplus consists of any benefit which he may be securing in any given transaction over and above that which he might secure by employing that money in other ways.

But now the question at once arises, what is it that distinguishes a producer from a consumer? In money economy it is customary to describe those persons as producers or sellers who offer goods in exchange for money, and those as consumers or buyers who offer money in exchange for goods. But this is a mere terminological convention and does not point to any essential difference. In the ultimate analysis the holder of money commonly described as the buyer is also a seller, when looked at from the point of view of those goods which he had sacrificed for procuring money. On the other hand, the producer is also a consumer of those goods which he secures in exchange for the resources employed. Indeed, in so far as the problem of economy turns out in the ultimate analysis to be a problem of maximising satisfaction out of given resources, every act of economic subjects may be represented as an act of exchange, and demand and supply or buying and selling as distinct processes may be said to have relevance only with reference to particular commodities.

In the present context, however, the distinction between producer's surplus and consumer's surplus will not relate to the character of the parties entering into exchange. We shall, on the other hand, draw a broader division between a surplus in terms of product emerging in the process of distribution of national income and a surplus gain in terms of utility arising out of exchange The former is more concrete and objective and has to be sought for within the so-called national dividend. The latter is essentially subjective, and refers, as we shall see, to an exchange in relation

to an individual economic subject, no matter what he is holding.[1]

2. In regard to producer's surplus, it is to be enquired if there is any category of income within the national dividend which can be described as an entity that is independent of cost of production. The answer to this enquiry has always been influenced by the type of problem economists had in view, and has importance not only form the theoretical point of view, showing as it does the gradual change in the economist's attitude towards what constitutes the 'economic problem',—but has also a very great practical significance.

Historically, a theory of distributional surplus,—a theory, for instance, that the 'social product' consists of two parts—cost and surplus—grew mainly out of an endeavour on the part of the early economists to evolve what may be called a theory of prosperity. Their preoccupation was principally to discover a criterion of the material prosperity of a society. Take the case of the Physiocrats. Behind all their treatment of what is today recognised to be the fundamentals of Economic Science—the problem of Value and Distribution—lurks an endeavour to show in what manner the landlord class might be made rich, for, as they thought, it was upon the riches of that class that the prosperity of a nation would depend. In the physiocratic system, 'product' is defined in physical terms, and 'cost' is described as the energy spent in production in terms of labour. Those enterprises are regarded as productive which yield a surplus, that is to say, which yield a return over and above the subsistence of persons engaged in them. Such surplus is said to arise only in agriculture where the co-operation of land yields a return

[1] This distinction will be brought out more clearly later. See Chapter VI, below.

which more than compensates for the wages of labour and the profits of stock. The physiocrats thus took land as the only productive factor and regarded the volume of rent as the index of the prosperity of a society.

In England, Malthus appears to have been most influenced by the physiocratic idea of surplus. Like the physiocrats, he considered rent as a bounty of nature, and regarded an increase of landlords' income as a sign of national prosperity. Many of Adam Smith's remarks on Rent also point to this, although, as we shall see, his general standpoint is in many respects different from that of the physiocrats. This view of surplus is derived from a conception of 'social return' and 'social cost' which runs in absolute terms.

On the side of cost, the early treatment was inadequate in view of its reliance on a crude subsistence theory of wages.[1] Wages were supposed to be rigidly fixed by the level of subsistence of labourers, so that labour cost could be estimated in absolute terms. To this the early economists were led by a rigid assumption that they held in regard to the growth of population. If wages tend to rise beyond the static level which is set by the subsistence of labourers, then, they argued, there would be an increase of population and the tendency would be checked. But in fact population growth does not admit of such a hard and

[1] A most rigid form of the subsistence theory of wages is to be found in Turgot and Adam Smith. "The wages of Workmen are limited to his subsistence by the competition among Workmen. He gets only his livelihood." Turgot, *Reflexions on the Formation and Distribution of Riches*, Economic Classics, edited by W. J. Ashley, sec. VI. Adam Smith also gives a similar theory of wages. "The demand for men" he says "like that of any other commodity, necessarily regulates the production of men", so that when "things are at a stand" the reward of labour is equal to its bare subsistence. It is only in an advancing society where the demand for labour runs faster than its supply that wages have a chance to rise beyond subsistence. Cf. *Wealth of Nations*, Bk. 1, Chap. VIII.

fast economic interpretation. The present-day tendency of a decreasing human fertility in almost all the civilised countries of the world shows beyond dispute that if there is any 'law' of population it is not merely economic.[1] There is therefore nothing like an absolute level of wages and no absolute measure of labour cost.

The misconception in regard to the 'product' of resources arose from a failure to distinguish between physical product and economic product. Now, so far as the physical product is concerned, there is hardly any clue to the division of the joint return among the co-operating agents. As Wieser points out, the question of such division is as absurd as to ask which part of the child is derived from the father and which part from the mother.[2] In Economics we are concerned with the value product of factors. What concerns us is not the division of a physical entity, but the *imputation* of the value return to different categories involved in production. And this imputation is based on the relative scarcity of factors in relation to demand. A particular agent may be productive in the physical sense; but may not be economically productive, just as a commodity may have value in use but no value in exchange. If land is abundant in supply and can be had for the mere asking, it is not economically productive, however necessary it may

[1] Modern historical researches tell us that even during the period which saw that great increase of population which set Malthus to discover an economic law of population, the increase was due to causes which have no direct bearing on the income of labourers. "Contrary to an opinion still widely held, the flood of life, which made Malthus and his generation speculate on the causes and cure of redundant population was due far more to life-saving, since the mid-eighteenth century, than to reckless procreation since the great inventions and the start of the Speenhamland policy of adjusting family receipts to the number of mouths to be fed." Clapham, *An Economic History of Modern Britain*, Book 1, Chapter 11, p.54.

[2] Wieser, The Austrian School and the Theory of Value, *Economic Journal*, Vol. 1.

be in the growing of corn. The entire value of corn is then distributed among other factors which are scarce, and nothing is imputed to land. This is the essential truth emphasized in Ricardo's theory of rent; and when it is suggested there that at the margin land yields no rent, the implication by no means is that land is not physically productive at the point—the physical co-operation of land is certainly essential at every stage of the process of cultivation—the implication is that land is not economically productive at a point where the entire product is exhausted in remunerating the other factors of production.

Now, under the influence of a purely materialist conception of 'productivity' and a conception of cost running in terms of the absolute subsistence of labour the physiocrats attributed the quality of productivity only to land. Wages, they argued, are meant to sustain the labourers,—to replace the energy spent in the process of production. It is land which is capable of producing a 'surplus' over and above this replacement cost, and the larger the volume of this surplus the greater, it was argued, would be the prosperity of a society.

Later on, however, in the hands of the socialists who regarded the entire social product as the creation of labour, this same 'surplus' appeared as an unearned income, measuring the degree of exploitation to which labour was subject under the capitalist system. There is indeed a close affinity between Marx's 'surplus value' and the *produit net* of the physiocrats; but, as is well-known, the contexts in which the concepts have been developed and used are fundamentally different. The theory of *produit net* related to a time when the new economic order—the capitalist system—had not yet developed and was meant, among other things, as a vindication of the landed aristocracy from whom it was supposed the society derived all its strength.

Marx's surplus value, on the other hand, was designed as a diatribe against the capitalist class—the creation of the new economic order. 'Surplus value' was supposed to be produced by labour but appropriated by the capitalist. It was thus presented as the source of a class antagonism, said to be inevitable in a capitalist economy.

But Marx's analysis of surplus value is also vitiated by an inadequate understanding of what constitutes the economic productivity of a factor. As we shall see, it has some relevance to a state of things in which the capitalists are enabled to exercise a monopoly power in the hiring of labour, but is entirely inappropriate for a competitive condition to which in fact it essentially relates.

Further, although there was a shift of emphasis in Marx's analysis from landlords' rent as the source of national strength and progress to the 'profits' of the propertied class as the source of antagonism between labour and capital, yet Marx also defined his surplus as a difference between an 'aggregate return' and 'aggregate labour cost'. Now, 'aggregate return' may have meaning only in a hypothetical single-commodity-community, but cannot be given any precise significance when it consists of heterogeneous goods. And in regard to 'cost', —in the first place, it does not consist merely of the labour employed,—in the second place, even the so-called labour cost cannot have any absolute measure, labour being of diverse quality and intensity. If, however, the two categories, 'aggregate return' and 'aggregate cost', are represented in terms of money, then obviously they are found to be two sides of the same thing. If the term 'aggregate social cost' can be used in any intelligible sense, it is the sum total of the prices of factors of production, and in that sense it coincides with the social money income, whatever the latter might be.

Theoretical Economics is concerned with the relative valuation of goods rather than with 'aggregates'. It aims above all to examine in what relations commodities stand to one another backed up by the preferences of economic subjects. And it is from these relative values of goods that factor prices are derived. The problem of Distribution is but an aspect of exchange value. It is a problem of the valuation of scarce factors. In its positive aspect, the theory of Value and Distribution in Economics is a neutral enquiry concerning the principles regulating the relative prices of goods and the imputation of value product to scarce means of production. It is a question of ratios rather than of an absolute measure. Costs and values are relative and not absolute concepts. It is true that an explanation of an inter-relationship of economic phenomena and the idea of a 'consequential adjustment' (order, equilibrium and so on) do involve a basic principle,—the principle of maximising. But the sense in which a 'maximum principle' may be said to lie at the basis of economic analysis is merely a formal one. The maximum principle that figures in the economist's system of equations has this simple, qualitative (as distinct from quantitative) implication that as consumer one prefers that good which satisfies a relatively more intense desire, and as producer one prefers that line which, other things being equal, promises to be relatively more productive. What is referred to is not an absolute quantity to be maximised, but rather a *force* which dominates human conduct in a world of scarcity. And this force works itself out through substitution at the margin.

This principle of relativity in economic valuation was indeed to a large extent recognised by earlier economists, particularly by Ricardo. Yet the bearing of the principle on the general field of economic science is more clearly shown in the works of Marginal Utility theorists.

From this point of view, the problem of distributional surplus leads us to an analysis of the relation between cost of production as 'displaced alternative' and the relative values of goods. We shall in particulr discuss the implication and the extent of validity of Ricardo's theory that 'rent does not enter into price'.

3. The theory of Consumer's Surplus relates to a gain in terms of utility arising from exchange, and is essentially a corollary of the 'maximum principle' operating in the sphere of consumer's choice. The idea of this gain from exchange is to be inferred from the bahaviour of the party concerned in the exchange and is therefore subjective. It is deduced from the fact that the consumer, in exercising his choice in the market, gives up that good which is lower on his scale of preference in exchange for one which stands higher. The 'utility' secured is greater than the 'utility' surrendered.

The idea is old, but the attempt to estimate it and to use it as an instrument for guiding economic policy is more recent. It was first suggested by Dupuit. But Dupuit's work was forgotten, and the concept was revived by Marshall, in whose system it plays an important part. Marshall realised the difficulties of estimating the surplus utility;[1] yet he suggested that within a narrow range the effects produced upon consumer's surplus by variations in the price of commodities could be *compared* when it is assumed that the commodities in question are 'independent', and, further, that the marginal utility of money that is sacrificed in the

[1] There is just a suggestion in the earlier editions of Marshall's *Principles* as if the technique of consumer's surplus (*net* satisfaction beyond consumer's costs) could be employed for measuring the aggregate economic welfare of the people. This is evident from the title of the chapter dealing with the concept which runs in the first two editions as 'The Measurement of the Utility of Wealth'. Later on, however, Marshall warns his readers against any such notion, and also alters the title

process of exchange is constant.[1] We shall examine in some detail the nature of consumer's surplus, and see how far it is an effective tool either for an understanding of the mechanism of exchange or for purposes of guiding economic policy.

4. Again, like all economic phenomena, the phenomenon of surplus has its static aspect and its dynamic aspect. Is it a phenomenon, it may be asked, which persists even when economic forces come to rest,—or is it one which is associated with economic change, itself opening out possibilities of shifts in the economic system? In the latter context we shall be particularly concerned with distributional surplus. Consumer's surplus, as we shall see, is a passive element in the economic process,—an 'economic epiphenomenon', as Johnson describes it.[2]

Dynamic surplus in the sphere of distribution is just a slice of the national dividend which, in the existing organisation of our economy, goes to the entrepreneur class upon whom rests the responsibility of fulfilling the economic function of adjusting the productive resources to the needs of the people. And it is the peculiarity of the capitalist economy where the motive of gain is the ruling force behind all economic activities that it is this surplus—or rather the expectation of it—that gives direction to the economic system. Thus Johnson describes dynamic surplus as 'economic surplus *par excellence*', for, it is this surplus that has 'economic poteney'. "To classify as surplus the incomes which at any given time promote change meets the claims of economic logic." [3]

[1] This assumption of constant marginal utility of money for ascertaining consumer's costs is comparable to the earlier notion of an 'invariable standard' in terms of labour used in connection with the measurement of cost of production.

[2] See his *Rent in Modern Economic Theory*, Chap. 1, pp. 10-12.

[3] *Ibid.*, p. 18.

Johnson's dynamic surplus, however, is a sort of 'frictional' income which appears in the wake of economic progress. Like a true Clarkian he contemplates the society as seeking to adjust itself to an ever-rising standard of economic life,—the surplus being an extra income earned during the process of adjustment by those entrepreneurs who initiate the change making for progress. This surplus has to be distinguished from the chance gains that arise on account of uncertainty. One is conceived essentially within a static framework, so that its appearance points to the movement of the economic system in an ascertainable direction. The other is an element in the causal process merely in the sense that its appearance reacts on the expectations of entrepreneurs and thus influences the direction of the economic system; but, as the nature of the reaction is not unique, the sort of movement that it calls forth is not also definite. Secondly, 'frictional surplus' of the Clark-Johnson model is something one may feel enthusiastic about. The entrepreneurs are cost-reducers,—raisers of the 'static standard'. If the cost reducing activities lead to an extra gain, it is not earned at the expense of any other class; rather it is calculated ultimately to improve the standard of the other classes as it gets absorbed in the general system. Chance gains, on the other hand—dynamic surplus in the strict sense—do not have any such claim. They indicate a less than 'optimum' output as set by the static standard and point to a feature which is a sort of a drag on the economic system. They are symptoms of social waste, —an evil, but perhaps a necessary evil. And the warrant for the entrepreneurial system of economic organisation,—the capitalist system, that is to say,—lies in so far as this evil is reduced to a minimum.

A few general remarks will be made on these problems in the concluding chapters where, in particular, attention

will be drawn to the distinction between pure dynamic surplus and what may be called 'frictional surplus'.

5. It has been noted that the answer to the enquiry concerning costs and surplus has been moulded from time to time by the particular bias economists had in regard to certain practical problems, which again were the product of particular environments. There have been two principal trends, one relating surplus to prosperity or welfare,[1] and the other relating it to class-antagonism. It is the latter that has played the more important part in the history of economic thought. Once it was recognised that 'product' could not be conceived except in value terms, the surplus of value over costs came to be regarded as the source of a clash of interest between different classes of income receivers. The Ricardian theory of rent stood for a critique of the landed interest, as showing a disharmony between the interest of landowners and that of the society; while Marx's theory of surplus value came as a critique of capitalism itself, appearing to have discovered a great contradiction that lay inherent in the capitalist system in view of a disharmony between the social interest and the interest of the capitalists themselves.

[1] In this respect the theory of consumer's surplus could perhaps be placed alongside of the Physiocratic theory of *produit net*. We have already referred to the suggestion in the earlier editions of Marshall's *Principles* about the net satisfaction beyond consumer's costs as measuring the aggregate economic welfare. Professor Pigou also prefaces his chapter on the measurement of National Dividend with the remark that the true measure of the aggregate economic welfare "involves the money figure that would be obtained by adding together the consumers' surpluses, as measured in money, derived from each several sort of commodity contained in the dividend." (*Economics of Welfare*, 3rd. edition, Chap. VI, p. 61.) This is not to say that either Marshall or Pigou failed to realise the difficulty of such measurement. This is merely to indicate the rather interesting family resemblance between the earlier objective notion of *produit net* and the later subjective notion of consumer's surplus.

It should be instructive to examine—as we have tried to do in the following pages—the possibility of such disharmonies in the light of modern economic analysis. It is sometimes alleged that modern economic theory, which starts from individual valuation and runs up to the imputation of the value product to the factors of production, evades the issue altogether, and that by so doing it loses much of the practical interest that economics ought to possess.[1] It is no doubt true that the modern theory of distribution, the marginal productivity theroy, as it is called, takes as a first approximation a world of perfect order where, under a system of perfect competition, factors combine among themselves as best they can and take back the proceeds of their joint effort according to a mechanical principle of marginal contribution,—the entire process being viewed as a condition of maximum economy relatively to given circumstances. The economic society thus presents itself as a big co-operative institution in which, whatever the immediate impulse moving them to action, the members do serve a social purpose, so that their relation turns out to be one of harmony rather than of discord,—of unity rather than of conflict.

But, of course, this is only a first approximation. In the actual economy there do prevail disorder and conflict

[1]See, e. g., Maurice Dobb, *Political Economy and Capitalism*. This important book came into my hands long after the present study had been completed. Otherwise it would come in for frequent reference in the text. The book is important in more ways than one. On a number of specific points concerning Ricardian economics—particularly in its practical aspect —and its affinity to the economics of Karl Marx it throws considerable light. In its general attitude, on the other hand,— particularly in its denunciation of the line of development of modern analysis,—it throws out a challenge and provokes arguing back. One may indeed fully sympathise with Dr. Dobb in his passionate urge to keep economic speculation close to earth, and yet dissociate oneself from his condemnation of the alleged shortcomings in this respect of modern economic theory.

which cry out for solution, and the economist cannot brush them aside. To explain them, however, he will look for a 'hitch' in the economic process. Now, this hitch is found to arise from three sources. There may be rigidities in the factor supply owing to specialisation of particular groups. Secondly, some factors in possession of monopoly may adopt policies inimical to social interest. And thirdly, there is the disturbing factor, uncertainty, that arises from imperfect knowledge. From the economic point of view, these are the objective basis of a conflict between general interest and the interest of particular factor groups. And the study of surplus, as distinguished from costs, may be associated with any of these phenomena.

Modern theory thus generalises, on an *analytical* rather than on a *classificatory* principle, the early investigations concerning rent and profit as sources of social disharmony.[1] Not only that. It does further make it possible for us to judge precisely to what extent these disharmonies are inevitable in the present or in any other institutional setting. And if that is so, it is surely well worth following it out.

[1] It thus appears to be beside the point to complain, as Dr. Dobb does, that "A principle which interprets value purely in terms of demand can define the productive 'contribution' of a person or a class only according to the value of what *eventuates*: it cannot define this contribution according to the activity or process in which the contribution *originates*." (*Op. cit.*, p. 30.)

DISTRIBUTIONAL SURPLUS

STATIC ASPECT

CHAPTER II

THE THEORY OF SURPLUS IN RELATION TO PROSPERITY

1. So long as Economics was studied in terms of greater or less of aggregate wealth there was naturally an endeavour to seek for a standard for measuring 'social cost' and 'social return', and to use the difference, if any, between the two as an index of a country's prosperity. This was the idea behind the physiocratic conception of surplus, —*produit net.* The famous *Tableau Economique* of Quesnay was designed to offer a formula for an ideal distribution of the annual income of a nation among different classes, and to examine and compare the accumulation of national wealth under such a process of distribution and the process that happened to exist at the time under all sorts of restrictions.[1] And Quesnay recommended perfect freedom of trade, because a 'natural order',—the condition of that ideal distribution,—would help the creation of the *produit net* which, according to him, was the basis of all progress. The society is composed of three classes, the Productive class, the Landlord class and the Sterile class. Those who are engaged in agriculture or other extractive industries

[1]This is not to say that it meant nothing beyond it. Indeed, strippped of certain naive assumptions in regard to the productivity function, the *Tableau Economique* could be shown to represent an algebraic equation which provides a description of a periodical flow of income through different parts of the economic system,—a description which is quite in line with any modern formulation. The point of emphasis is rather how in it the conception of the ideal and the conception of the natural stand mixed up, and how the analysis was primarily an analysis of prosperity.

form the productive class. It is they who produce a surplus over and above what they require for their subsistence. This surplus or *produit net* they pay to the landlord. The manufacturers and the traders do not produce any such surplus. They form the sterile class and are the wage earners of the productive class. The annual income of a nation is thus identical with the *produit net* from land, and the problem of distribution resolves itself into the mode of circulation of this *produit net*, —an examination of the manner in which the expenditure of the *produit net* by the landlords influences the economic condition of a nation.

The theory that the volume of surplus produce that goes to the landlords and the manner in which it is spent shape the economic position of a nation was already in Cantillon,[1] who is thus sometimes considered to be a precursor of the physiocratic system. As soon as a society of men is formed, the ownership of cultivable land goes to some few persons, and the farmers and labourers, after they have provided for their own living out of the proceeds of the land, transfer the surplus to these proprietors.[2] It is these proprietors who are independent, and all other classes, hired labourers or undertakers, live at their expense. These dependent classes 'live from day to day', and their earning and spending are passive elements in society. It is the 'Surplus du Produit' which is the active factor making for social progress. The volume of this surplus conditions the development of cities. The disposal of the surplus, the fancies, fashions and modes of living of landowners, give direction to the nature of commodities produced. Even the number of population in the society depends upon

[1] This was also the theory of William Petty; cf. *A Treatise on Taxes and Contributions* (incorporated in Monroe's *Early Economic Thought*).

[2] "The overplus of the Land (*surplus du produit*) is at the disposition of the Owner". *Essai sur la Nature du Commerce en General* (edited by Henry Higgs), p. 7.

the way the surplus is disposed of, that is to say, upon the purpose to which land is turned.[1] So far the *Essai* clearly anticipates the physiocratic-system, and it is believed that Quensay, the founder of the system, drew considerably from Cantillon.[2] Like the Physiocrats, Cantillon gives a supreme role in production to land, the 'Fountain-Head', as he calls it, from which flow all the means that support the inhabitants. Like them, too, he defines the income of landlords as a *part* of the produce of land,—a surplus over and above the cost of maintenance of labourers and the undertaking of farmers. He suggests furthermore that it is this surplus which is the source of all progress. But there are passages in the *Essai* which, taken independently, demonstrate that it has a more modern ring about it. In analysing the problem of value, for example, Cantillon suggests that the 'intrinsic value' (normal price in modern terminology) of a commodity, as distinct from its market price, corresponds to "the quantity of Land and of Labour entering into its production, having regard to the fertility or produce of the Land and to the quality of the Labour",[3] —a proposition which the later economists might do well

[1] Here, however, Cantillon introduces a qualifying clause the importance of which had not been sufficiently recognised for a long time since he wrote. After having stated the manner in which the size of population is conditioned by the subsistence offered by the proprietors he adds that there is also the psychological element to be considered. "When I said that the Proprietors of Land might multiply the population as far as the Land would support them, I assumed that most men desire nothing better than to marry if they are set in a position to keep their Families in the same style as they are content to live themselves. That is, if a Man is satisfied with the produce of an acre and a half of Land he will marry if he is sure of having to keep his Family in the same way. But if he is only satisfied with the produce of 5 to 10 acres he will be in no hurry to marry unless he thinks he can bring up his Family in the same manner." *Essai*, p. 77.

[2] See Jevons, *Richard Cantillon and The Nationality of Political Economy*, incorporated in Higgs' edition of the *Essai*, pp. 354-355.

[3] *Essai*, p. 29.

to remember. Again, in spite of the supreme importance that he attaches to land, he mentions quite correctly the relation of interdependence between the elements of production. "The Land belongs to the Proprietors but would be useless to them if it were not cultivated. The more labour is expended on it, other things being equal, the more it produces, and the more its products are worked up, other things being equal, the more value they have as merchandise. Hence the Proprietors have need of the inhabitants as these have of the Proprietors."[1] Furthermore—and this is perhaps the most important part of his teaching so far as pure theory of Economics is concerned—Cantillon does not identify national dividend with the 'surplus du produit', but takes the entire flow of income in a society as the object of distribution, and seeks also to discover a 'par' between the value of land and labour.[2] The national dividend consists of three parts which he designates as the 'three rents' of the farmer. The 'principal or true rent' is that part which the farmer pays to the proprietor. A second part he spends for his maintenance and the maintenance of the labourers employed to cultivate the farm, and a third is a compensation for the risk of his undertaking. Here we have the rudiments of a more scientific theory of distribution; and although Cantillon did not solve the problem, he stated it more correctly than the economists who immediately followed him.

[1] *Op. cit.*, p. 47. These statements, perhaps more than the one quoted by Jevons in the essay above referred to, justify his contention that Cantillon's *Essai* "holds the balance betwween the elements of production more evenly than any subsequent treatise." See pp. 342-343.

[2] In this he is indebted to Petty. Petty recognised the necessity of translating one factor in terms of another so as to discover the relative costs of production of goods in the circumstances of joint production. In fact, following Schumpeter, we might regard the modern marginal productivity analysis as an attempt at a solution of 'Petty's Problem'. Cf. *The Theory of Economic Development*, p. 23.

There are obvious crudities in the physiocratic view of production and distribution which have long since been exposed and upon which it is needless to expatiate. What is fundamentally wrong there is, as is well known, that the physiocrats failed to detect the economic significance of production, that they made a confusion between matter and utility,—a confusion which of course persisted in some form or other in subsequent writers, too, until it was cleared up by J. B. Say. If, as is emphasized by Say, production means creation of utility, not a creation of matter, the dividend in which the economist has got to be interested relates to the entire flow of utilities or benefits that come into being and are valued in the market, and the problem of distribution turns out to be one of an apportionment of the flow of income or dividend thus defined among the factors that co-operate to bring it into being. It is, in other words, a problem of the valuation of factors of production, or 'goods of higher orders', to use a modern terminology. So far is the modern view of production removed from the physiocratic view that even factors of production have ceased to be considered as technical entities and have come to be regarded as bundles of 'productive services'.

Now, it is clear that the physiocratic conception of surplus as something *social* and material, offered as it were as a gift to society by nature and consisting of the extra yield from agriculture and other extractive industries beyond the cost of subsistence of the labourers employed in them, is derived chiefly from two assumptions. In the first place, it is assumed that production means creation of new physical matter. This assumption apart from its intrinsic weakness, begging as it does the question whether at all it is physically possible to create new matter, has just been seen not to have any significance from the economic point of view. And secondly, it is assumed that cost is to be

measured by the subsistence of labour. The second assumption is the basis of what is known as the Iron Law of Wages. It rests on the ill-founded notion that population has always a tendency to adjust itself to the means of subsistence,—ill-founded indeed, for it neither has an empirical validity, nor can it be deduced from any *a priori* reasoning, so many extra-economic factors influence the growth of numbers.

2. It is clear, therefore, that any judgment on the economic condition of a nation on the basis of a social surplus of the sort conceived of by the physiocrats would be not only meaningless, it would be misleading. But the concept of *produit net* was used in another very important context, too. If the surplus produce of land is the fund which creates the prosperity of a nation by giving activity to other professions, by stimulating commerce and industry and by favouring population, it is also the only fund on which the state can draw for meeting public expenditure. It is the only taxable income of a society. All other earnings are mere returns to labour employed or advances made. They are expenses of production, and form no net revenue. If any of these earnings are at all taxed, the tax is ultimately shifted to the landlord's surplus. "There exists no truly disposable revenue in a state except the net produce of lands".[1] Convenience and expediency therefore require that state expenditure should be made out of a direct appropriation of a part of the landlord's income. A single, direct tax on this 'disposable' surplus would leave the yearly production of the national produce intact, and would therefore be economical. This is the physiocratic theory of '*Impot Unique*'. So great was the influence of this doctrine that in more places than one attempts were made

[1] Turgot, *Reflections on the Formation and Distribution of Riches*; Economic Classics (ed. Ashley), sec. XCIX.

at the time to put it into practice. The experience, however, did not prove successful.[1] Later, by the end of the last century the authority of the physiocrats was invoked by Henry George in what is known as the 'single tax' campaign.

A proper analysis of the theory of *impot unique* shows that it was essentially a problem in incidence of taxation. The physiocrats were in fact seeking for a portion of the social product which the state might touch without doing damage to productivity. They tried to evolve a tax system in which the tax payer might be himself the tax bearer, a system which would eliminate the costly process of 'shifting'. And with that notion of 'net product' before them they were naturally driven to consider the income of the landowning class as *the* source to draw upon for defraying state expenditure. It was as if asking this particular class to do the political job gratis.[2] But while it is true that they introduced a really interesting problem for the economist, their analysis of the problem was wrong and their solution was necessarily crude. This will be discussed later.

3. Since the time of the physiocrats the idea of surplus has been looming large in economic discussions. There has always been a tendency to associate it with land rent, as the physiocrats did, so much so that today whatever appears in economic analysis as a surplus is designated as rent. But the early notion of surplus was vitiated by its identification with a social physical product. We have seen how the physiocrats conceived it like that and held the volume of surplus or landlord's income as an index of a country's prosperity. A somewhat similar attitude appears

[1] See Higgs, *The Physiocrats*, p. 86.

[2] For a fuller comprehension of the real analytical spirit behind the theory of *impot unique*, see Einaudi, The Physiocratic Theory of Taxation, *Economic Essays in Honour of Gustav Cassel*.

to have characterised Adam Smith's theory of rent, too; although his general system differs in fundamental respects from that of the physiocrats. Adam Smith had also a theory of prosperity which contrasted social cost with social surplus, and one does notice traces of physiocratic influence in some of his remarks relating to the theory of rent. "Land, in almost any situation, produces a greater quantity of food than what is sufficient to maintain all the labour necessary for bringing it to market in the most liberal way in which that labour is ever maintained. The surplus is too always more than sufficient to replace the stock which employed that labour, together with its profits. Something therefore always remains for a rent to the landlord." [1] This surplus yield of land stands out in sharp contrast with cost which represents the subsistence of labour and is something definite. With every improvement in the productive powers of labour the surplus has a tendency to increase. "Every increase in the real wealth of the society, every increase in the quantity of useful labour employed within it, tends ..... to raise the real rent of land." [2] The interest of the landlord is thus 'strictly and inseparably connected with the general interest of the society'.

All this sounds physiocratic indeed. But if that is so, how are we to take the statement with which Smith opens up his enquiry,— that "the annual labour of every nation is the fund which originally supplies it with all the necessaries and conveniences of life which it annually consumes", —a statement which is somewhat of a reaction against the physiocratic enthusiasm for land? There is indeed much confusion in Smith's theory of rent which we shall have occasion to notice later on. But some explanation of the

[1] *The Wealth of Nations* (edited by Cannan), Bk. 1, Chap. XI, p. 147.

[2] *Op. cit.*., Bk. 1. Chap. XI, p. 247.

above paradox is found in his chapter on *Wages of Labour*. In the original state of affairs, when land is not appropriated and capital is not accumulated, it is labour alone which enjoys all that it creates by exploiting its environment. Nobody else is there to share its fruits. In such a state the annual produce of a nation is all the work of labour and destined to be shared as wages among the labourers themselves. But this peaceful enjoyment of the fruits of their toil by the labourers cannot continue long. With the progress of society land becomes private property and capital begins to accumulate in the hands of a few. New social classes spring up and they begin to lay claim upon the produce of labour. Thus emerge rent and profits,—deductions from the gross produce of labour employed upon land. But why at all should the labourers submit to this deduction? Because they have to,—because, once they are dispossessed of land, and once they get to depend for their subsistence during the interval between sowing and reaping the harvest upon advances made by the capitalists, these masters and landowners would not agree to part with the services of their property unless some compensation were made to them. And as men multiply in proportion to the means of subsistence, and as therefore food is always in demand, some compensation can always be extorted out of the labourers, leaving them with the bare means of livelihood. Rent and profits are thus both a surplus beyond the cost of subsistence of labour, and bear no proportion to 'the quantity, the hardship, or the ingenuity' of those who own them. And this surplus is a 'monopoly price', a charge on the fruits of labour.

The entire volume of this surplus, however, is not to be taken as an index of prosperity. The prosperity of a society depends on the aggregate real wealth that it can secure through the employment of labour. Now, whereas

an increase in real wealth is attended with high rent, it makes for low profits by favouring a larger accumulation of capital.[1] The same cause which raises rent lowers profits. The interest of the landowner and the interest of the capitalist are differently related to the general interest of the society. It ought to be noted that when Adam Smith says that high rent goes with general prosperity he does not mean, as the physiocrats used to do, that high rent is *itself* a factor making for social progress. High rent is taken by Smith to be a symptom of increasing prosperity, the two phenomena being co-effects of the same cause, —an increase in the productive powers of labour employed upon land.

Malthus' allegiance to the physiocratic theory of rent is even more conspicuous. Rent, Malthus says, is the surplus produce of land which arises from a 'bountiful gift of Providence'. The emergence of rent is "a clear indication of a most inestimable quality in the soil which God has bestowed on man—the quality of being able to maintain more persons than are necessary to work it."[2] It is not a monopoly price enjoyed by the landowners at the expense of the labourers, but a natural return of land which manifests itself in the early periods of society in the form of extraordinarily high profits and extraordinarily high wages, but emerges as a separate category when, with the progress of society, a larger accumulation of capital lowers profits and a tendency of population to grow beyond the means of subsistence lowers wages. This surplus produce of land again—Malthus echoes with the physiocrats—is at the root of all that makes for the prosperity of a society. It is the source of all power and enjoyment. If the surplus is small, society is ill-provided with conveniences and luxuries and

[1] Cf. Turgot, *Reflections*, sec. LXXXIX, where a similar suggestion is made.

[2] *On Rent and on Foreign Corn*, p. 16.

leisure, while, if it is large, 'manufactures, foreign luxuries, arts, letters and leisure may abound'.

4. The problem of distribution was yet primarily a problem of the disposal of a surplus conceived as a collective physical entity and marked off from the wages of labour. The problem of production was yet a problem of what causes led to the prosperity of a nation. Had Adam Smith and Malthus turned their attention from the theory of prosperity, which led them nowhere, to the theory of economic valuation they would soon have discovered that the reason why with the progress of society rent came out as a separate category of income was that as population grew land became scarce in relation to demand; it became an economic good and its use came to command a price in the market. They would also have discovered that being a phenomenon resulting from scarcity it could not be regarded as an index of prosperity.

Both Adam Smith and Malthus seem to have been feeling their way to an explanation of the phenomenon of rent, but they just missed it, and failure to catch the simple truth drove them into absurd conclusions. How absurd they could be is amply illustrated by the following sentence which occurs in Malthus' essay on *Rent*: "The cause of the high price of the necessaries of life above the cost of production, is to be found in their abundance rather than their scarcity".[1] As the context would show, it was a conclusion to which Malthus was inevitably led by his initial error of considering rent as the bounty of nature, or, to be more precise, of not distinguishing between physical productivity and economic productivity. The same criticism applies to all those remarks of Adam Smith which suggest that high rent is the result of high productive powers of labour employed upon land. Adam Smith, however, might have

[1] *Op. cit.*, p. 13.

at the back of his mind some sort of an idea of scarcity when he described rent as a monopoly price.[1] But the term 'monopoly' was certainly misplaced in this context. Scarcity is not the only attribute of monopoly. Were it so, every economic good would be a monopoly good. Moreover, Adam Smith cannot be said to have seen through the logical possibilities that the scarcity explanation of rent would signify.

It is the achievement of West[2] and Ricardo to have brought out this fundamental principle governing the rent of land,—the principle of scarcity. The apparently paradoxical statement of Ricardo that "the labour of nature is paid not because she does much, but because she does little", contains this essential truth that the ultimate cause of the value of land is the scarcity of the services offered by it, for it is this characteristic that makes it an economic good. This discovery of the cause of rent marks a decisive shift in the central preoccupation of economists away from the problem of 'riches' to the problem of 'value'. It was a shift—significant from the point of view of the present study—from an enquiry into an aggregate volume of surplus produce to serve as an index of prosperity to a more neutral enquiry whether there is any factor in the productive process whose income does not exercise any influence on the relative value of goods, but is itself derived from it. Ricardo's answer to the enquiry is that there is one such factor—

[1] It looks rather plausible, this confusion between monopoly and limitation of supply; for, even Senior who did so much by way of emphasizing and bringing out clearly the aspect of scarcity in economic goods fell a victim to this confusion when it came to the question of the scarcity of gifts of nature, and described rent as a monopoly price. See below, Chap. V.

[2] The credit of discovering the cause of rent goes to West, for it was he who first recognised the significance of scarcity and of 'margin', although it was Ricardo who fitted them into the general body of economic analysis. See West, *The Application of Capital to Land* (1915).

that comprising the 'original and indestructible powers of the soil'. The payment made for these powers only he calls rent, and he examines how the relative value of goods is determined at a margin where no rent is paid, and how rent is itself a 'differential surplus' depending upon the value of goods and arising from the intra-marginal units of labour and capital. True Ricardo's rent theory is, as we shall see, in many particulars not altogether free from criticism. But the concentration on an 'intensive' margin that the analysis implied had a profound significance in that by pointing to the very important distinction between physical productivity and economic productivity it opened the way to a clearer understanding of the economic problem of distribution.

CHAPTER III

A CRITICAL EXAMINATION OF THE THEORY OF SURPLUS VALUE AND UNEARNED INCOME

1. In some respects analogous to the physiocratic theory of *produit net*, albeit contemplated with an essentially different outlook and a different objective, is Marx's theory of surplus. The conception of surplus developed by Marx runs in terms of an aggregate quantity. It is an excess of aggregate return over aggregate cost; and this aggregate cost is represented by the cost of maintenance of the total labour power of a society, "the socially necessary labour", to use a Marxian phrase. Again the hypothesis of a crude subsistence theory of wages on which the earlier conception of surplus was based.[1]

But the parallelism ends here. Unlike the physiocrats Marx considers 'form' as important as 'matter' in economic calculation, and concentrates on value product rather than on physical product. It is thus a theory of surplus *value* and not one of net physical product, the content of the former being wider than that of the latter. The process that creates surplus value is not necessarily 'productive' in the physiocratic sense.[2] 'Surplus value' is the genus of which 'produit net' is a species. In the second

[1] "The value of labour power is the value of the means of subsistence necessary for the maintenance of the owner of labour power". *Capital* (Everyman's Library), p. 155.

[2] "The physiocrats declare that agricultural work is the only kind of productive work; for this, they consider, is the only kind of work that produces surplus value. For them surplus value exists only in the form of land rent." Marx, *op. cit.*, p. 553; see also p. 185n.

place,—and this is the cardinal point—the origin of surplus is attributed by Marx to an excess product of labour and not to an excess product of land. Surplus value is all the creation of labour but is appropriated by the owners of instruments, – land[1] and machinery. Whereas the physiocrats considered surplus as a gift of nature for which no labour is spent, Marx takes it as a product of labour for which no wage is paid. With the physiocrats the volume of surplus is the measure of the prosperity of a society. With Marx it is a measure of the degree of exploitation of labour by capital.

Marx starts with two general principles in formulating his theory of surplus value. It is assumed in the first place that labour is the sole source of value. Commodities in so far as they are objects of exchange are imagined by Marx as so much 'congealed labour', as mere phenomenal manifestations of a mass of human labour power. Labour is that common element which imparts the attribute of exchangeability to goods that are apparently so diverse in form and character. The value of goods therefore is determined by the quantity of labour embodied in them, that is to say, upon the quantity of labour worked up in producing them. "What determines the magnitude of value of a use value is the amount of socially necessary labour time it contains",[2] and "it is the magnitude of the value of a commodity that determines its exchange relations."[3] Secondly, it is argued that the exchange value of labour, its wage, is determined by the cost of subsistence. This second principle, as Marx maintains, is just a corollary of the other one. It is an extension of the labour theory of value to the special case of the commodity, labour. It is not the natural outcome of any inherent

[1] "The most general instrument of labour is the earth." *Ibid.* p. 173.
[2] *Op. cit.*, p. 8. By a 'use value' Marx here means a commodity.
[3] *Ibid*, p. 35.

tendency of population to adjust itself to the means of subsistence, but rather a necessary concomitant of the capitalist system, —a system in which labour as such is bought and sold. In so far as labour itself takes the character of an exchangeable commodity, its value depends like that of any other commodity upon the quantity of labour required to produce it, —in other words, upon the quantity of labour involved in the production of just the amount of commodities necessary for its reproduction and maintenance.

Now, under the circumstances of a capitalist system the quantity of labour used up in the production of commodities is greater than the quantity of labour that is necessary for the maintenance of the total labour power. The 'use value' of labour, as Marx calls it, is greater than its exchange value. Hence arises a surplus which is distributed as rent, interest and profit. Not that labour works out of nothing. But the instruments with which it co-operates only carry forward the value that they originally possess, and the simple additions of these pre-existing values are no cause of surplus value. Labour alone is instrumental in producing something which is more than its market value. Labour alone, of all factors, is *creative*.

But why is it that under capitalist production labour produces more than it consumes? Why is the surplus value that labour creates appropriated by somebody else? Because it is the characteristic of capitalist production that workers are enabled to work in a group, and through co-operation and appropriate division they can produce more than they could do had they worked in insolation. A 'combined working day' produces larger quantities of commodities than an equivalent sum of 'isolated working days'. Capitalism brings out economies of division of labour and economies of co-operation. At the same time, however, it robs the workers of their independent status

and of the power to use their labour and to enjoy the fruits thereof freely. In the first place, they cannot own the instruments on which the process of production depends,—land and machinery. The workers have no capital. In the second place, in order that the workers who come to co-operate might work in harmony, the process of production is to be under the control of an authority. And the function of this co-ordination, this bringing the individual units of labour power into an organised, co-operative whole, devolves upon one who has command of capital. It is thus an essential feature of capitalist production that workers are subordinated to capital. The worker sells labour and is paid according to its exchange value which is equal to the cost of its subsistence. But the labour itself or the use value that it produces becomes the property of the capitalist. It is then the business of the capitalist to make the labour that he purchases yield more than it costs, to squeeze out of it surplus value. It is the surplus value that motivates capitalist production. And the capitalist secures it by one or both of two processes: extension of the period of work beyond what yields the workers' subsistence, and intensification of the productive capacity of workers through improved organisation. In the one case the workers are made to put forth *more* labour than what is required for their subsistence. In the other case *less* labour is made to yield the required subsistence.[1] The volume of surplus then is a function of the number of workers, hours of work and the technique of production, and it measures the extent to which labour is exploited by capital.

2. This in short is Marx's theory of surplus value, hailed by socialists as forming a scientific basis of their attack on capitalism. It is supposed to have demonstrated

[1] The former process bears what Marx calls 'absolute surplus value', and the latter process bears 'relative surplus value'.

logically how capitalism contains in it an inherent evil, the evil of exploitation of labour by capital.

But a moment's reflection shows that the so-called logic is built on very shaky foundations. The labour value theory on which Marx bases his analysis of surplus value is taken over from Adam Smith's philosophical account of value,[1] —a theory of value which Smith conceived as appropriate for a primitive state of society in which land is not appropriated and capital is not accumulated,—in which, therefore, labour is the only economically significant factor in the productive process. Now, of course, even in such a society an adequate explanation of the phenomenon of exchange value—and it is exchange value that concerns us in Economics—is not provided by an unaided labour theory. The units of labour employed may be—and indeed are—of different quality and different intensity, and this contingency makes the situation less easy to handle. Marx avoids the difficulty by referring to an average. The labour which determines value does not represent this or that individual unit; it is a homogeneous mass, each unit being counted in relation to a 'social average labour power'. But how are we to translate the heterogeneous units as they are into an abstract homogeneous mass? Where is the common denominator,—the average—in terms of which to express the different units of labour? To find out such a denominator is itself a problem which cannot be solved independently of the wages of labour determined in the market and depending in their turn upon the value of goods.[2] Yet, with all this, as a study of the historical process of economic evolution Smith's labour

[1] For a discussion of Adam Smith's philosophical and empirical accounts of the theory of value, see below, Chapter IV.

[2] For a general criticism of Marx's labour value theory see Wicksteed, *The Marxian Theory of Value*, reprinted from *Today* in *The Commonsense of Political Economy* (ed. Robbins), Vol. II, pp. 705-724.

value theory can be given some meaning. When land could be had for the mere asking and when capital had not yet begun to function the entire produce of a society would belong to labour as a class. Again, as a basis of abstract reasoning such a starting point is helpful. An isolated individual working under primitive conditions with enough of land and to spare will take his labour as the standard for estimating cost and will put the relative value of goods that he acquires accordingly. But it is only a starting point. It only brings into relief the complex elements that have to be taken into account as the economy passes from the stage of simple production to that of joint production. When production becomes joint, when, that is to say, more than one factor enter into the productive process, exchange value depends, among other things, upon the relative proportion in which the factors are used in the production of different commodities and the rate at which they are remunerated.

Marx saw through the contradiction. His analysis of exchange value as it is in the market in fact brought him to a point where he discovered that the amount of labour embodied in commodities did *not* determine their exchange relations,—that the prices of commodities had a tendency to deviate from the 'magnitude of value' as he defined it. He discovered, for example, that if in the production of different commodities varying proportions of 'constant capital' and 'varying capital'[1] are employed, then these commodities cannot exchange in proportion to the labour attaching to them if the rate of profit in different industries is to be equalised as it must be. Where the commodity requires capital of higher composition, that is to say,

[1] These terms are used by Marx. They correspond respectively to the more familiar terms Fixed Capital and Circulating Capital. That portion of capital out of which wages are paid is variable capital and the other portion is constant capital. *Op. cit.*, Vol. III, Part II, Chapters VIII-X.

where the percentage of constant capital is greater, price is higher than 'value', and where the percentage of constant capital is less, price is lower than 'value'.[1] But instead of this impasse leading to a rejection of the labour theory and the absolute value concept, it only led him to the conviction that individual deviations of prices from value were unimportant in themselves; and he went at pains to prove how the reciprocal fluctuations of prices round value would cancel each other, so that taking all industries together the total of the prices of commodities would be equal to 'the sum of their values'. But what earthly meaning can be given to an expression such as 'the sum of values'? As Allyn Young has observed, the summation of economic values is akin in principle to the task of determining the weight of the solar system. Both are equally impossible of achievement. Indeed, the entire analysis of Marx in this connection is worked out against a background of mysticism which, whatever its merits might be otherwise, has no place in a study which aspires to be at all scientific.

Again, the subsistence theory of wages, another of those posts on which the surplus value theory is erected, is merely an extension of the hypothesis of labour as the measure of value. As has been pointed out above, it is only an application of this general proposition regarding value to the particular case of the exchange value of labour. It may therefore be immediately argued that both stand or fall together, and with them the theory of surplus value. Yet it is profitable to examine the theory a little more closely to see just where its positive deficiency lies.

The capitalist form of commodity circulation, the form that is associated with surplus value, is, as Marx describes

[1] For a full discussion of the contradiction involved in Marx's labour cost theory of value and a most perverse attempt at a solution of the contradiction that he made, see Böhm-Bawerk, *Karl Marx, passim.*

it, a transference of money *via* commodities to money,—M—C—M. The capitalist spends money on acquiring commodities of one description, transforms them into commodities of another description and sells them for money again. But the money thus acquired differs in magnitude from the money originally invested. The process thus becomes M—C—M', where M' is greater than M by, say, Δ M. This Δ M emerges on account of less money being spent on one of the commodities in the process of circulation than its 'use value'. And that commodity is labour. ΔM is attributed to labour, so that from the standpoint of the capitalist it appears as a surplus. In attributing this increment of value, ΔM, to labour alone, Marx assumes that the other instruments that move in the circuit, although they pass through different forms, carry a constant value during the whole process. But do they really? Does it not often happen that even without the least intervention of labour a commodity gets higher value added to it as time passes? The example of a bottle of wine or a tree is familiar to all readers of Ricardo and Jevons. Time is an element in the value creating process. In the economic sphere, as in any other sphere, they also serve who only stand and wait. Marx's idea that labour produces more than is paid for, and that this phenomenon explains the increment of value alluded to, betrays a complete ignorance of the principle governing the economic productivity of a factor under joint production. When production is joint, how are we to discover what amount of labour would be necessary for the workers' subsistence, and what specific amount of commodities is produced by the labour actually employed? Marx himself argues that it is by virtue of co-operation that the productive process yields an excess of commodities over what is required for the workers' subsistence. Take away those facilities for co-operation and more labour will be

point. In any system these earnings would emerge. Whether, of course, society ought to value such services in the way it does, or whether these persons ought to appropriate the so-called surpluses, are different questions into which we need not enter here.

3. A failure to detect the economic significance of the productivity of a factor also characterises the theory which suggests that rent is an 'unearned' income and is a peculiarly suitable object of taxation. We have seen that the physiocrats considered the income of landlords as the only taxable revenue of a society. Writing in this connection, Turgot observes, "The class of proprietors,.... not being bound by the need of subsistence to a particular labour, can be employed for the general needs of the society.... either by a personal service, or by payment of a part of their revenue with which the state or the society may engage men to discharge these functions."[1] A similar but yet stronger sentiment is expressed by Adam Smith when he casts a fling at the landlord class that they 'reap where they never sowed'. The proprietors of land, he says, "are the only one of the three orders whose revenue costs them neither labour nor care but comes to them, as it were, of its own accord, and independent of any plan or project of their own".[2] These remarks seem to have been the basis of an attack on private property in land by many even of those who are otherwise counted as 'orthodox' economists. John Stuart Mill, too conservative to suggest any disturbance on the institution of private property favoured an appropriation by the state of what he thought to be the accretions of income attaching to land as a necessary result of progress. "The ordinary progress of a society which increases in wealth, is at all times tending to augment the

[1] *Op. cit.*, Sec. XCVI.

[2] *Op. cit.*, Bk. I, Chap. XI, P. 248.

The sort of surplus here referred to may arise out of monopoly in the commodity market or out of monopsony in the factor market as well.[1] Marx's theory may be said to relate to the latter case. And there his generalisation is surely unwarranted. It is not certainly true that capitalism—defined in a normal way as a system in which the motive of profit works as a spur to economic activity — contains in it anything that makes for monopoly in general.[3] Much less is it likely, so long as labour is not 'specific',[4] that there should be monopsony in respect of the 'buying' of labour. Although a producer might enjoy a monopoly in a commodity market, the employees would yet be in a position to turn to other occupations, so that in regard to the buying of labour there would be competition among employers.

Moreover, as has been noted, the interpretation is purely formal. Except in the case of what may be termed 'institutional monopoly', — those kinds of monopoly, that is to say, that are created and fostered by legal institutions, monopolies can hardly be called exploitative. Monopolies are no doubt productive, and so long as they arise out of natural scarcity there is no competitive criterion with reference to which we could ascertain the extent of surplus earnings of these monopoly agents. Their capital value, on the other hand, is to be reckoned with reference to the value that is imputed to their scarce services. 'Inventors and Money-makers' of exceptional genius are cases in

[1] Cf. Mrs. Robinson, *The Economics of Imperfect Competition*, Chaps. 25, 26.

[2] It should be noted in this connection that the assumption of perfect competition is not that an individual entrepreneur *cannot* influence the market price. It would be indeed absurd to assume that. The assumption is that he *does not expect* that a variation in his own output would influence the market price. Stated in this way the problem of Imperfect Competition does not seem to be as important as has been made out in recent years.

[3] About specificity more will be said in Chapter V.

point. In any system these earnings would emerge. Whether, of course, society ought to value such services in the way it does, or whether these persons ought to appropriate the so-called surpluses, are different questions into which we need not enter here.

3. A failure to detect the economic significance of the productivity of a factor also characterises the theory which suggests that rent is an 'unearned' income and is a peculiarly suitable object of taxation. We have seen that the physiocrats considered the income of landlords as the only taxable revenue of a society. Writing in this connection, Turgot observes, "The class of proprietors,.... not being bound by the need of subsistence to a particular labour, can be employed for the general needs of the society.... either by a personal service, or by payment of a part of their revenue with which the state or the society may engage men to discharge these functions."[1] A similar but yet stronger sentiment is expressed by Adam Smith when he casts a fling at the landlord class that they 'reap where they never sowed'.. The proprietors of land, he says, "are the only one of the three orders whose revenue costs them neither labour nor care but comes to them, as it were, of its own accord, and independent of any plan or project of their own".[2] These remarks seem to have been the basis of an attack on private property in land by many even of those who are otherwise counted as 'orthodox' economists. John Stuart Mill, too conservative to suggest any disturbance on the institution of private property favoured an appropriation by the state of what he thought to be the accretions of income attaching to land as a necessary result of progress. "The ordinary progress of a society which increases in wealth, is at all times tending to augment the

[1] *Op. cit.*, Sec. XCVI.

[2] *Op. cit.*, Bk. I, Chap. XI, P. 248.

income of landlords; to give them both a greater amount and greater proportion of the wealth of the community, independently of any trouble or outlay incurred by themselves. They grow richer, as it were, in their sleep, without working, risking or economising."[1] These increments are an 'unearned appendage' to the wealth of landlords which ought to be transferred to the state and spent for the general benefit of the society.

But where to fix up the standard rent increment beyond which could be considered as 'unearned'? How to abstract the 'unearned' from the 'earned' income? Does, moreover, an increment of rent necessarily follow in the wake of economic progress? If not, would the state give compensation for 'undeserved *decrement*' as it would confiscate 'unearned increment'? Faced with these difficulties Sidgwick proposed a wholesale nationalisation of land. But as a sudden action was bound to be attended with serious disturbances in the economic system, the process, he urged, should be slow and gradual. The policy of land nationalisation received yet stronger support in America and on the continent towards the end of the last century. In America in particular there was a great campaign for it under the leadership of Henry George whose book, *Progress and Poverty*, attained at its time a degree of popularity comparable to that of Marx's *Capital*.

Although at the present moment it has only an academic interest and is not seriously considered except as a part of the socialist programme, yet certain difficulties of land nationalisation may be pointed out here.

One thing should be noted before we proceed to discuss the question. Do the terms 'earned' and 'unearned' used in this context refer to an income as it stands in relation to the corresponding factor of production, or to the particular

[1] *Principles of Political Economy*, Bk. V, Chap. II, Sec. 5.

persons who happen to possess the factor in question? For, it is one thing to enquire concerning the imputation of product to a factor, it is another and a different question whether, judged by the standard of equity, persons owning particular factors of production deserve what they earn,—this dichotomy in distribution theory arising particularly from the fact that in regard to non-human factors of production the owner is not to be identified with the factor as such. Now, the latter problem is one which hardly lends itself to an objective analysis. It is with an eye to this aspect of the problem that John Stuart Mill made the statement which has since become famous, namely, that distribution, unlike production, obeys no rigid law.[1] The standard of equity varies and is reflected in social laws and institutions, the laws and institutions again determining what part of the social product each person within their jurisdiction would have the right to enjoy. What Russia will decry as an 'unearned income' may perhaps have ungrudging sanction in a more or less individualistic society like that of England.

Not so, however, with factorial distribution, an impersonal affair of imputation, which transcends all institutional limits. From the point of view of imputation surely land is on all fours with other factors with which it co-operates and surely rent is not an unearned income so long as land is paid for because—and to the extent that—it is productive. No matter who owns it, so long as land is scarce in relation to the demand, its income is a natural and an inevitable condition of the economic process.[2] The

[1] "The laws and condition of production of wealth, partake of the character of physical truths... It is not so with the Distribution of Wealth. That is a matter of human institution only." *Principles*, Bk. II, Chap. I, Sec. I.

[2] For an elaborate discussion on how even under a socialist regime rent of land would emerge as a natural outcome of the economic process, see Wieser, *Natural Value;* also Cassel, *The Theory of Social Economy*, Vol. I.

implication of Ricardo's theory of rent, an implication which it is very important to realise, is that land, and for the matter of that any factor, becomes economically productive only when demand for its use extends so far that application of other resources to it begins to yield diminishing returns. And when this stage is reached, not only is it not uneconomical to pay for its use according to productivity, it is uneconomical not to pay; for such payment keeps the investment of resources on it on the *economic margin*.[1] Any other arrangement would be inconsistent with the maximum principle.

Now, for fixing upon that structure of normal prices for different categories of land which secures the most economical distribution as between different productive channels, ownership of land is essential. And it is one of the virtues of competition and free contract that they tend automatically to effect such distribution. If the ownership is transferred to the state, then the state would become a thorough monopolist and there would be no guarantee that the rent charged should obey the marginal productivity principle,[2] nor would it be easy for the state, even if it chose, to discover the competitive level of rent.

One might argue, however, that the state would choose to apply the competitive instead of the monopoly principle in the hiring out of land, and might further suggest that

[1] On this point see Knight, Fallacies in the Interpretation of Social Cost, *Quarterly Journal of Economics*, Aug., 1924; reprinted in *The Ethics of Competition;* also his *Risk, Uncertainty and Profit*, Chap. IV.

[2] This observation may appear paradoxical; for, one of the points made out by agrarian socialists is that land itself is a monopoly factor and that state ownership is calculated just to break the private monopoly. But the argument is based upon a confusion between ecarcity and monopoly. Land is scarce but is not necessarily a monopoly. There is nothing in the nature of land by virtue of which an individual landowner can have a control over the price of land any more than an individual wheat-grower can have a control over the price of wheat.

if it is a question of mere transference of land from private to public ownership with all other competitive forces intact, then the difficulty of discovering the competitive structure of rent would not be serious, for the state would simply place all available lands on the market and offer them to the highest bidders. But even then the position is less simple than at first sight it appears. We cannot, for all we know, assume that there is a clear-cut supply of land in the market with all possibilities of further exploration absent. And moreover we cannot ignore the problem of finding a criterion for weighing the alternative possibilities of investment of capital on land development and investment of capital in other lines.

Apart from these difficulties it may be further asked,—how is the process of nationalisation to be carried out in practice? If through purchase, then the purchase price would certainly include the future anticipated income of land, in which case there is just a chance that the inequality of income should be accentuated rather than mitigated. The bugbear that we were out to fight would only appear in another form. Perhaps the very persons who were hitherto receiving rent from the cultivators would now receive a perpetuity from the state. And who knows if the move would not be in the undesired direction?

For, as we have already hinted, there is no guarantee that with the progress of society rent has a tendency to rise. The course of a factor income absolute or relative to other incomes is affected by economic progress in a variety of ways. It is less simple than was contemplated by the early economists. An increase in the number of population, for example, is not necessarily followed by a fall in the wages of labour. The character of the general productivity function, the condition of supply of other factors and the elasticity of substitution for them may be such that even

the relative share of labour—not to speak of its absolute share—should arise on account of an increase of population. Nor is the course of rent so certain. There is no evidence to show why it should not "crook and turn upon itself in many a backward streaming curve".

4. It is said that rent arises out of a conjuncture of circumstances, land having no supply price. Turgot, it will be remembered, placed rent as a 'disposable' income as distinct from interest on the ground that "the price one pays for the purchase of land does not contribute in any way to the revenue it produces."[1] Historically this contention may be regarded as the genesis of what later developed into the proposition, 'Rent does not enter into price'. We shall discuss this theory in greater detail in subsequent chapters. It need only be pointed out here that in a sense all economic values are the result of a conjuncture of circumstances. In an exchange economy commodities derive their significance from environmental conditions. The reason why for one's wares one gets a particular price and no less is that his neighbour has got a specific demand for them. It is true that the productivity of land does not depend on a price being paid for it; but so is it with any factor of production. It is not because a price is paid for it that a factor is productive; a price is paid for a factor because it is productive.

Can it not be argued, however, that the remuneration of a factor exceeds that which is necessary to evoke a particular supply of the services offered by it? Does it not often happen that the circumstances of the market confer upon a factor a price which is too high absolutely, so that an appropriation of a part of that price through taxation would not disturb the supply of the given factor that was forthcoming or of the commodities that it was producing? It is a surplus of this sort that Hobson in particular tried

[1] *Op. cit.*, XCVI.

to emphasize.[1] Every factor of production has got a certain maintenance price which has got to be paid if the industrial system is to be kept going. The productive process involves some loss of energy, some depletion of resources which have got constantly to be replenished. The payments that are necessary for keeping intact this energy through which the industrial system functions constitute what Hobson calls the 'cost of maintenance'. In other words, they are stationary supply prices of factors of production. Any income beyond this supply price is, according to Hobson, a surplus. It is an excess of income over the cost of maintaining a stationary economy. A portion of this surplus, however, is essential for the growth of the industrial system, and is therefore productive. It consists of those extra remunerations which have got to be paid in order to evoke an additional supply of productive agencies appropriate for a growing system. These remunerations are considered as the 'cost of growth'. If the surplus extends beyond this, it ceases to be productive, being a superfluous payment not necessary either to support or to stimulate industrial activities. It becomes an 'unearned income'. All 'economic rent' of land and other natural resources, all interest in excess of that required to give a stimulus to saving adequate for the growth of the industrial system, all profits and wages beyond what is just necessary to evoke the sufficient use of the relevant factors belong to this category. These payments are not only superfluous because the industrial system would take its course independently of them; they are harmful, being a source of industrial waste. They lead to over-investment and under-consumption, and are responsible for those oscillations of industry and employment which are familiar occurrences in the modern economic system. An appropriation by the state of this 'unproductive

[1] Cf. his *The Industrial System*, Chap. IV; *The Science of Wealth*, Chap. V.

surplus' is conducive to the smooth working of the system, and far from reducing the volume of industrial output increases it by compelling those who live upon it to work more.

At first sight the argument appears to be plausible. We can certainly conceive of a factor having a supply curve parallel to the vertical axis, at any rate over a certain range, so that, so far as that particular range is concerned, a reduction of its price through taxation would not react on its total supply. We can even conceive of a factor with a supply curve which is 'backward-rising', in which case such reduction of price would be followed by an increased supply. But does it mean that their services are at any particular point unproductive? The scarcity of a factor from which arises economic productivity has to be considered not absolutely but relatively to demand. When the pressure of demand for a factor increases and the supply does not respond to it, it only means that the less important uses are surrendered and the factor becomes more productive of utility. There is thus no supply price of any factor in the absolute sense. The supply price of a factor with reference to any particular industry is derived from the importance of the next best possible use to which it might be applied.

In the same way again Hobson's explanation of industrial oscillations is surely misleading. These oscillations have little to do with the emergence of the alleged 'unproductive surplus'. In the first place, there is no reason why if one earns more than what just keeps him at his job, the extra income should not be *spent* towards an increase of the standard of living. In the second place, even if it is *saved*, it cannot occasion any trouble so long as there are investment opportunities open.[1]

But on all this more later on.[2]

[1] Cf. Robbins, *Consumption and the Trade Cycle*, *Economica*, November, 1932. Also E. F. Durbin, *Purchasing Power and Trade Depression*.

[2] See below, Chap. X.

CHAPTER IV

LABOUR COST AND THE DIFFERENTIAL SURPLUS THEORY OF RENT

1. We saw earlier that it would be purely arbitrary to slice the 'social income' into two parts, cost and surplus, as suggested in the writings of the older economists,—that much of the confusion leading to such division arose from a failure to distinguish between physical productivity and economic productivity. A more interesting problem for the economist is to enquire concerning the relation between factor incomes and the value of goods of the first order. From this point of view, that income category is to be looked upon as a surplus which is not a condition of the value of goods but merely its effect. Ricardo's doctrine that 'rent is not a component part of the price of commodities' has this implication, that rent—defined as a payment for 'the use of the original and indestructible powers of the soil'—is a differential surplus and is not *causal* in relation to price.

2. Even before Ricardo economists were very often in the habit of looking upon rent as the effect of price and not its cause. We have already referred to the suggestion of Turgot in regard to the distinction between rent and interest. This standpoint is also taken by Hume in a letter which he wrote to Adam Smith on the publication of the *Wealth of Nations*. "I cannot think", he says, "that the rent of farms makes any part of the price of the produce, but that the price is determined altogether by the quantity and the demand."[1]

[1] Quoted in Hirst's *Adam Smith* (English Men of Letters), p.164.

Turning to Adam Smith, one is apt to feel puzzled over a number of contradictory statements that he makes on the relation between rent and price. In the sections dealing with *Value* it is stated that in an advanced society, where land is appropriated, rent forms a component part of the price of commodities. "As soon as the land of any country has all become private property, the landlords.....demand a rent even for its natural produce", so that "in the price of the greater part of commodities the rent of land comes ...to constitute the third source of value."[1] Again, "the natural price itself varies with the natural rate of each of its component parts, wages, profit and rent."[2] In the chapter on *Rent*, on the other hand, a different view is taken. "Rent", it is stated there, "enters into the composition of the price of commodities in a different way from wages and profit. High or low wages and profit are the causes of high or low price; high or low rent is the effect of it. It is because high or low wages and profit must be paid in order to bring a particular commodity to market, that its price is high or low. But it is because its price is high or low; a great deal more, or very little more, or no more, than what is sufficient to pay those wages and profit, that it affords a high rent, or low rent, or no rent at all."[3]

This contradiction in Smith's statements is generally explained with reference to his twofold interpretation of the problem of value, which, following Von Wieser, it has been customary to call the 'philosophical' account and the 'empirical' account.[4] Adam Smith promises to investigate the principles regulating the 'exchangeable value' of goods; and yet he is anxious to discover a relation between this ex-

[1] *Wealth of Nations* (edited by Cannan), p.51.

[2] *Ibid.* p.65.

[3] *Ibid.* p.147.

[4] See Wieser, *Natural Value*, (English translation, edited by Smart), pp. XXVII-XXVIII.

changeable value with what he calls the 'real price' of goods. By 'real price' he means the quantity of labour, the 'toil and trouble' which is required to produce a commodity. It is a purely metaphysical concept. He imagines in this context a primitive, 'rude' state of society in which land is not appropriated and capital is not accumulated. In that imaginary society, value of goods is regulated by the quantity of labour that is bestowed on their production; there is no asymmetry between the 'real price' and 'exchangeable value' of goods. "If among a nation of hunters .....it usually costs twice the labour to kill a beaver which it does to kill a deer, one beaver should naturally exchange for or be worth two deer."[1] In the actual society in which he lived, however, this simple explanation was not sufficient. In the 'improved society' where lands are appropriated and capital is accumulated, the 'exchangeable value' and 'real price' are two different things. In the circumstances of an improved society the exchangeable value is regulated not by the quantity of labour involved in the production of goods but by the remuneration of the different factors that enter into their production. Thus the three component parts of price are wages, profit and rent. This is the empirical account of value, a theory of value which is relevant to the actual state of things in an advanced community.

Now, in the imaginary society of Adam Smith in which land is not appropriated, the question of land rent does not arise at all, for land in such circumstances is not an economic good, and no rent need be paid for its use. On the other hand, when rent emerges in the economic system it is considered to be an element of price. What does Adam Smith, then, mean by isolating rent to be the effect of price and not its cause? The explanation seems to be that while giving his empirical account of value he is looking at the question

[1] *Op. cit.*, p. 49.

from the point of view of an entrepreneur, and is asking at what price it pays the entrepreneur to sell his commodity,—what, in other words, are the elements that make up the entrepreneur's supply price.[1] In the other context, on the contrary, he is concerned with the evolution of rent as a social phenomenon; and there he is still haunted by his 'real price' concept. In the hunting stage referred to above when labour is the only factor involved in the production of goods the real price is an index of exchangeable value, because the whole of what is produced is absorbed by labourers, wages being the only category of income. In the advanced community a commodity exchanges for more than the quantity of labour embodied in it (labour having its material counterpart again in wages), and to the extent that it does it yields rent and profit. But whereas profit depends upon the supply and demand of 'stock', rent is merely an outcome of increasing population, and consequently a transfer income.

And, of course, this is the general attitude maintained in the *Wealth of Nations*. There is an undercurrent of a philosophy of value pervading the entire system of Adam Smith.[2] In the enunciation of the theory of rent he was more actuated by what is generally known as the real cost

[1] This is clear in a passage in the chapter on *The Component Parts of Price* which runs thus: "....a third circumstance must be taken into consideration, the rent of land, and the commodity must commonly purchase, command or exchange for, an additional quantity of labour, in order to enable *the person who brings it to market* to pay this rent". *Ibid.* p. 52n. (Italics are mine). This passage occurs in the first edition and is dropped in later editions.

[2] It ought to be noted that although Adam Smith, generally speaking, is actuated by a philosophy of value, he makes a number of suggestive remarks about the alternative uses of land in which there is a clear emphasis on the phenomenon of the market and on the fact that the rent which a piece of land could secure elsewhere influences the price of the crop for which it is actually used. See, e.g. *Wealth of Nations*, Bk. 1, Chap. XI Part I.

concept. Rent is a component part of price, but is no ingredient of cost in the sense of real cost ; for, as he says, labour of nature 'costs no expence'.[1]

This becomes particularly clear from his attitude towards profit. Adam Smith's description of the transition from the primitive to an improved society suggests that profit has also to be conceived of as a transfer income co-ordinate with rent. And, in fact, as was pointed out earlier, there are indications that he considers profit, too, as a surplus. In the chapter on *The Wages of Labour* he emphasizes a monopoly aspect of profit resulting from a combination of employers (the owners of stock), and throughout his treatment of profit the advantage of breaking this monopoly and a consequent lowering of profit is demonstrated. Yet it is clear that profit as such is not considered by him as co-ordinate with rent so far as its relation with price is concerned. In the passage quoted above, profit goes with wages as a determinant of price unlike rent.

The explanation of this apparent inconsistency is that having no clear idea about the distinction between monopoly and scarcity, Adam Smith could not draw a sharp line between normal profit and monopoly revenue. His treatment of profit in general relates to normal profit. It is a recompense for sacrifice in the form of waiting and risk (not that these terms are used by him), and is therefore different from rent. Whereas the landlords 'reap where they never sowed', the capitalist earns profit against

[1] "In agriculture too nature labours along with man; and though her labour costs no expence, its produce has its value, as well as that of the most expensive workman." *Op. cit.* Bk. II, Chap. V, p. 343. These remarks occur in a passage which more than any other in the book suggests an affiliation to the physiocratic theory of productivity and indicates, as was pointed out earlier, how Adam Smith had an insufficient appreciation of the scarcity explanation of the value of land.

advances made towards the maintenance of the workers till the completion of their work and in view of the fact that he 'hazards his stock' in the enterprise.[1]

Herein is in a nebulous state a concept of 'real cost' which comes into sharp relief in the writings of the later classical economists. It appears in a clear and more explicit form in Senior and is then followed up by Cairnes in whom it finds its strongest adherent. This aspect of cost is developed in terms of an absolute disutility,—a positive sacrifice undergone in the productive process. The essential ingredients of cost consist in the psychological feelings of discomfort or 'pain' involved in the use of productive factors. The objective counterpart of these feelings in the form of money outlays is a derivative phenomenon. Emphasis is given on 'sacrifice' rather than on 'reward'. The use of 'natural agents' involves no cost in this sense. Any income that accrues to such agents is a surplus which emerges because—and to the extent that—price deviates from real cost. The relevance of this conception of cost to the explanation of the relative value of goods will be examined in the next chapter where it will be shown that it bears a wider sociological rather than a strictly economic import.

3. But, is Ricardo's enunciation of the differential surplus theory of rent with its corollary that 'rent does not enter into price' by itself imbued with any philosophical bias?

It is often thought that Ricardo's was a pure labour theory of value,—that all that Ricardo did in this context was to attempt an extension of Adam Smith's philosophical account of value to the complex conditions of an advanced society. The introductory remarks in Whitaker's interesting study of the history of the labour theory of value lend colour to this idea. In fact Whitaker seems to deny to Ricardo any claim towards making an advancement of the

[1] Cf. *op. cit.* Bk.I, Chap.VI, p.51; Chap. VIII, p. 67.

empirical account of value.[1] Similarly, commenting on 'Ricardo's Attempt to Revive the Pure Labour Theory', Cannan observes, "His (Ricardo's) heart clung to the pure labour theory, but his candid brain showed him that it was weak,"[2]

If these statements are meant to show, as they apparently are, that there is anything like a disutility cost concept predominant in Ricardo's analysis of value, then they do not seem to present a true view of Ricardo's position. It is true that Ricardo in starting his discussion on value quotes with approval Adam Smith's real price doctrine, yet the manner in which he applies it in his own scheme is devoid of any metaphysical significance.

Unlike Adam Smith and Malthus, Ricardo does not concern himself with an explanation of the causes of the prosperity of a nation. As he points out in the preface to his *Principles*, the main problem of Economics is to determine the laws regulating the distribution of income among different factors of production.[3] This is a perfectly neutral enquiry and at once rules out the problem of a contrast between real cost and income. Whereas Adam Smith's discussion of value is subsidiary to his more fundamental problem of the production of wealth,—division of labour

[1] "Ricardo contributed very little to the advancement of the empirical account as such. The direct line of descent of this doctrine is traceable from Smith's *Wealth of Nations*, through the *Principles* of Malthus and J. S. Mill, to Marshall. Neither Ricardo nor Cairnes can be considered to stand in the line." *History and Criticism of the Labour Theory of Value*, p. 14. In the chapter dealing with Ricardo's theory of value, however, this extreme view is not maintained.

[2] *A Review of Economic Theory*, p. 177.

[3] He emphasizes it again in a letter to Malthus. "Political Economy, you think, is an enquiry into the nature and causes of wealth; I think it should rather be called an enquiry into the laws that determine the division of the produce of industry amongst the classes which concur in its formation." *Letters of Ricardo to Malthus* (edited by Bonar), p. 175.

which influences such production working itself out through the mechanism of the market, in Ricardo the problem of value assumes its significance because a variation of the relative prices of goods influences the proportion in which income is distributed into wages, profit and rent. Real price in the sense of an absolute value of goods plays the least part in Ricardo's system. Ricardo fixes his eyes on the relative price of goods and seeks to discover the causes leading to *variations* thereof.[1] And if he places particular emphasis on the element of labour as a regulator of relative prices, it is not out of any metaphysical predilection directed to a measurement of the sacrifice in the form of pain involved in the production of goods and to bringing out a contrast between the real price as dictated by pain-cost and the actual price in the market, but merely to indicate that relative prices are affected by causes which act on a particular commodity more than by causes acting on commodities in general. It is not an emphasis on 'toil and trouble' as distinct from wages; it is an emphasis on the amount of wages to be paid on account of more or less labour being employed on a particular commodity as distinct from the *rate of wages*. If the proportion in which labour and capital are used in the production of goods is the same, an increase or decrease in the rate of wages would affect all commodities alike and would therefore be neutral towards relative prices. In such circumstances, however, a variation in the quantity of labour required to produce a particular commodity, that is to say, a variation in the amount of wages spent on it, other things remaining the same, does lead to a variation in its value in relation to other commodities. But, of course, this proportion is not the same for all commo-

[1] This is a point which it is very important to bear in mind. We fail to understand properly Ricardo's analysis of value unless we remember that it is a study in Comparative Statics rather than in the Theory of Equilibrium.

dities, and if it were the same, any of the factors could equally be shown to be a regulator of relative prices. Ricardo was pretty clear about it. "If", he writes to Malthus, "all things required precisely the same quantities of capital and labour, and for the same length of time to produce them, any one of them would be an accurate measure of the rest; but this is not the case; the conditions admit of infinite variety, and therefore whichever we choose it can only be an approximation to truth, and we are bound to give good reasons for preferring it."[1]

Ricardo thus emphasizes that in so far as in the production of goods unequal proportion of fixed and circulating capital is employed, and in so far as it is returned to its employer with unequal rapidity, the rate of wages and profits, too, does operate in the regulation of relative prices. If the rate of wages rises "all commodities which are produced by very valuable machinery, or in valuable buildings, or which require a great length of time before they can be brought to market, would fall in relative value, while all those which were chiefly produced by labour, or which would be speedily brought to market would rise in relative value."[2] And conversely for a rise in the rate of profits.

These emendations to the original proposition with which Ricardo starts demonstrate beyond doubt that his analysis of value is far from being guided by a labour philosophy. They serve further to show how erroneous is the notion prevalent in certain quarters that Marx's theory of value which we have discussed earlier bears an affinity to Ricardo's theory. In fact such a notion rests on a gross misunderstanding of Ricardo. Whereas Ricardo is quite

[1] *Letters of Ricardo to Malthus*, edited by Bonar, p. 222.

[2] *Principles of Political Economy and Taxation* (Gonner's Edition), p.29. See in this connection Ricardo's Letters to McCulloch (edited by Hollander), Nos. XIV, XV, XXII, XXXII, XXXIX, XLIV.

clear about his conception of value as a ratio, Marx, as we have seen, works with a conception of value as a magnitude. Whereas Ricardo's sole preoccupation is to explain value as a market phenomenon, Marx assumes value to be a metaphysical entity and asks if there is any correspondence between the value thus conceived and the value as it is in the market. The two studies belong altogether to different planes.

In support of his contention that a philosophical account dominates Ricardo's theory of value,—that it is all an attempt "to force the empirical principles, or the 'facts' of entrepreneur's cost, to fit the labour philosophy" Whitaker mentions Ricardo's persistent use of the term 'labour' instead of 'wages' as denoting cost of production and as regulating natural price. He argues as follows:

"If Ricardo were working with merely an empirical account of value, and were not embarrassed by an uncertain philosophy of value, how would he ever come to speak of the cost of production, which determines value, as consisting of 'labour and profits'! He should say 'wages and profits'. Again, how would he be led to commence his chapter on 'Natural and Market Price', by the assertion that the market price of a commodity can deviate temporarily from 'its natural price, or the quantity of labour which it has cost!' No empirical theory would lead to the statement that the normal price toward which the competition forces actual prices is a *quantity of labour*. This natural price is Adam Smith's 'philosophical' natural price or 'first price', namely, labour".[1]

Now, it is true that certain expressions of Ricardo, if taken without due recognition of the general background of his system, might be considered to have a metaphysical complexion. Yet it seems pretty obvious that what Ricardo means by labour in the contexts referred to by Whitaker is

[1] *Op. cit.*, p. 131, footnote.

'wages' and not 'toil and trouble'. Had he meant the latter, the expression 'labour and profits' would be ridiculous indeed. 'Toil and trouble' cannot certainly be taken as co-ordinate with profits, for one relates to 'states of consciousness' and the other relates to the phenomenon of the market. In fact no such psychological category is involved in Ricardo's use of the term 'labour'. Whenever Ricardo speaks of cost of production he means expenses in money terms, as a few quotations given below will show:

"The real value of a commodity I think means the same as its cost of production" and "a commodity is at its just or natural value, when it repays by its price, at the usual rate all the expenses that have been bestowed upon it, from the first to last to produce and bring to market." [1]

Again; "By cost of production I invariably mean wages and profits."[2] The following passage is more explicit on the point in question: "Mr. Malthus accuses me of confounding the very important distinction between cost and value. If by cost Mr. Malthus means the wages paid for labour, I do not confound cost and value, because I do not say that a commodity the labour on which cost £1000 will therefore sell for £1000; it may sell for £1100, £1200, or £1500,—but I say it will sell for the same as another commodity the labour on which also cost £1000; that is to say, that commodities will be valuable in proportion to the quantity of labour expended on them. If by cost Mr. Malthus means cost of production, he must include profits, as well as labour." [3]

Here is a passage in which the expression 'labour and profits' occurs as standing for cost of production; but can it be

[1] *Notes on Malthus* (edited by Hollander and Gregory), p. 15.

[2] *Ibid*, p.18.

[3] *Ibid*, p. 14. Cf. also *Principles*, Chap. I, Sec. VI, the concluding paragraph and the footnote attached to it.

doubted but that the term 'labour' is used to signify money wages and not anything like a psychological discomfort?

This passage, moreover, shows particularly clearly that what Ricardo was concerned with is a ratio of exchange and not an absolute value of goods. Ricardo's 'natural price' has thus an entirely different significance from the 'philosophical' natural price of Adam Smith. It stands for a normal price ratio as between goods that tends to be established in the market. It is a mistake to confuse it with an investigation that relates to discovering a measure for the value of a good by itself.'[1] It is indeed far from Ricardo's mind to identify natural price with a quantity of labour. Contrary to what Whitaker's comments would suggest, in the chapter on *Natural and Market Price* Ricardo does not use 'natural price' as synonymous with labour cost, but only refers to the comparative quantity of labour used in the production of commodities as the predominant element regulating their relative prices in the long run.[2] Whereas even in his empirical account of value Adam Smith takes labour as an invariable standard for measuring the cost of production of goods, Ricardo makes it quite clear that neither labour nor, for the matter of that, any commodity would serve the purpose of an invariable standard, and also that a commodity could have as many values as there are other commodities to which it is related. "There has never been", he writes to McCulloch, "and.....never will be any perfect measure of value."[3] It is true that the gold index that he

[1] "To confound the two ideas would be much the same thing as to overlook the distinction between the thermometer and the fire." J. S. Mill, *Principles*, Bk. III, Chap. XV.

[2] I have searched in vain in Ricardo's writings for a statement like 'natural price, or the quantity of labour which it has cost.' I do not know where Whitaker quotes it from.

[3] *Letters to McCulloch* (edited by Hollander), p. 173. Malthus also recognises that the idea of a specific measure of value is inconsistent with

adopts as serving approximately, as he says, his own purposes of analysis[1] is not free from criticism, as it does contradict the strict implication of the principle of value as a ratio,[2] yet, from the point of view which we have been emphasizing here, it will be well to remember that it postulates quite a different kind of hypothesis to that which Adam Smith's labour command measure does. It is an index and not a measure,—an index of the variation of relative values. If gold and other commodities are produced under the same circumstances, so far as the proportion in which different factors (labour, fixed and circulating capital etc.) are employed in them is concerned, then a variation in wage does not affect the prices of commodities, because the relation between them and the medium, gold, remains unaltered.[3] If, however, owing to the introduction of certain facilities, one of the commodities requires less labour for its production, then, other things remaining the same, it will fall in value in terms of gold, and in the contrary case it will rise in value. He gives no doubt an undue emphasis on labour as a factor of production, but his analysis excludes any assumption concerning the subjective cost of labour. It examines the expenses of production *from the point of view of an industry*. It does not seek to measure them in terms of 'sacrifice' *from the point of view of the labourer*.

Adam Smith's definition of 'exchangeable value'. Yet he gives his own definition of value as "the power of purchasing, arising from intrinsic causes", and fixes upon labour as the measure of that power. Cf. his *Principles of Political Economy* (2nd edition), Chap. II, secs. 5 & 6.

[1] E.g., to study the relation between wages and profits on the one hand, and the price of goods on the other.

[2] Cf. Samuel Bailey, *A Critical Dissertation on the Nature and Measure of Values* (1825); London School Reprint (1931).

[3] It should be noted that one of the aims behind Ricardo's analysis of value was to show that higher wages would not necessarily mean higher prices.

4. Ricardo's main preoccupation was to evolve a principle which regulates the proportion in which the annual produce of a country is distributed between the factors that bring it into being, and he tried to show how a variation in the relative prices of goods affects this distribution. Ricardo wrote at a time when England was still under the burden of laws restricting importation of corn and when, on account of increasing pressure of population on land on the one hand and mechanical inventions facilitating manufacture on the other, the price of corn was rising as against other goods, and rent of land was going high. His problem was to provide a theoretical explanation of this phenomenon.

Two questions thus presented themselves before Ricardo: the variation of relative prices of goods, and the progress of Rent, Wages, and Profits. Now, taking 'utility' of goods for granted he concentrated on the forces operating on the supply side. With the progress of society there is an increasing demand for food and an increasing pressure on land for growing corn. But owing to the scarcity of land and owing to the operation of diminishing returns, more labour has to be employed to raise an equal amount of corn as cultivation progresses. The price of corn is determined at a margin where it is raised at the highest cost in labour. The excess of this marginal cost over average cost is rent per unit of product. The margin shifts downwards as population increases, and corn rises, giving landlords higher rent at the expense of the owners of capital.

Ricardo defines rent as the income from "the original and indestructible powers of the soil", that is to say, from a factor the supply of which is absolutely fixed and invariable. As it is 'original', the supply cannot be increased by human industry. As it is 'indestructible', the supply is not diminished through decay. The value of such a factor depends solely on demand and is derived from the price of goods

for which it is used. The causal relation runs one way—from commodity price to factor price. Hence his oft-quoted dictum,—"Corn is not high because a rent is paid, but a rent is paid because corn is high." The volume of rent depends upon the position of the margin, or in other words, upon the extent of demand in relation to its given supply.

This demonstration of Ricardo marks the beginning of marginal analysis in Value and Distribution. It introduces the idea of 'differential' into economic reasoning, and points to the element of scarcity as the foundation of the value of land. The modern Imputation theory of Distribution is but an application of the Ricardian theory of Rent into yet wider fields.

But Ricardo's observations on the rent price relationship and on the progress of factor incomes have only a limited validity, and have often given a handle to false conclusions. On the relation between rent and price more will be said in the next chapter. It may only be pointed out here that the validity of Ricardo's dictum depends largely on how we define an industry. If by 'corn' we mean agricultural commodities in general, and if we assume that manufactured goods require so small a proportion of the total supply of land that its demand from that quarter might be neglected, or that 'urban land', to use a comprehensive term for all lands required for non-agricultural purposes, is not in effective competition with agricultural land, then Ricardo's dictum does hold true. And in fact Ricardo had some such idea in mind. His theory is directly influenced by the contemporary economic condition of England, e.g., a continued rise in the price of agricultural produce as against manufacture.

But this does not by any means provide a precise explanation of economic equilibrium. The assumptions above noted are arbitrary. The lumping together of all agri-

cultural produce under one comprehensive category, 'corn', has no analytical justification. It conceals from our view all the tendencies to adjustment that have to be taken into account so far as different crops themselves are concerned. Moreover, even the quantity of land allotted in a society to agricultural operations as such is not a rigidly fixed quantity, but is determined with reference to the possibilities of its use in other directions, including non-pecuniary uses. As such, rent is as much a factor influencing relative prices as wages and profits are. The same allowance that Ricardo makes for these latter elements has to be made for rent, too. If, for example, by the fiat of an authority there is effected a wholesale reduction of rent, the measure will affect different commodities differently. The normal check to the application of land to relatively less profitable channels will slacken, and the 'margin of transference'[1] will alter. It is not true, as Ricardo would imagine, that the effect of an act of surrender on the part of landlords of the whole of their rent would be merely that the farmers would 'live like gentlemen'. The effect would be a reshuffling of land as between different uses in favour of those goods which require a relatively larger proportion of land for their production.

As regards the progress of rent, as we have already seen, no generalisation can be made. Ricardo's observations are based on the assumption of a very low elasticity of substitution between land on the one hand and labour and capital on the other. It is only as a special case that he allows for improvements in agriculture leading to a fall in rent. Under Ricardo's normal assumptions, if, from the social point of view, the supply of land is fixed, and population is increasing, it is true that the share of the national income that goes to the landlord would have a tendency to increase. But there

[1] Cf. Henderson, *Supply and Demand*, pp. 94-98.

cannot be any general presumption to that effect, as recent experience would suggest. With the growth of 'manufactured' land competing with 'natural' land (to use Prof. Cassel's apt expressions) and with the tendency observable in many countries towards a stationary population, the situation appears to be the reverse of that which Ricardo had envisaged. Far from the pressure of population on available land being the case, the problem in many places might tend to be rather a pressure of means of subsistence on population!

Ricardo's formulation of the differential surplus theory of rent is in one fundamental respect thoroughly misleading. Where he talks of the existence of differences in the fertility of different soils and derives a surplus on the basis of an 'extensive margin', he ignores the fact that these lands, not being homogeneous, cannot form members of the same factor group,—that they cannot be brought under the same supply curve unless they are capable of being expressed in terms of some common standard of efficiency. And if they are so expressed any surplus of the sort contemplated by Ricardo loses its significance. It is only a higher price paid for a factor having superior productive capacity, —and in terms of 'efficiency unit'—a larger total value paid for the use of a larger quantity of the factor. In principle there is no distinguishing between the superior productive power of a piece of land and the superior natural ability of a person. Reliance on an extensive margin for the explanation of rent, apart from having this analytical defect, moreover conceals the fact that the heterogeneous pieces of land, although not forming the same factor group, have within them a high degree of substitutability, so that the existence of inferior lands is a check to the rise of rent. Ricardo's way of putting things has just a risk of being understood to mean as though resort to inferior lands is a cause of

high rent, whereas in fact it is a force by which the rise the rent of superior lands is counteracted.

But although this aspect of the differential surplus theory has received prominence in economic text-books, the more important part of Ricardo's analysis is that which relates to a surplus based on the 'intensive margin',—a margin reached as a result of the continued application of doses of labour and capital on a given supply of land,—a margin at which all that is produced is absorbed by the variable factors and nothing is left for land. Referring to the question whether there is in any country a no-rent land, Ricardo observes that it is immaterial and that his theory does not fall through "if there be any capital employed... on land which yields only the return of stock with its ordinary profits, whether it be employed on old or on new land".[1] The cultivator with a given supply of land will 'dose' capital (capital standing for all factors other than land) up to the margin of profitableness, that is to say, up to the point at which the marginal product is equal to the value of capital. And because, with diminishing returns operating, the productivity of earlier doses is higher, a surplus arises which is to be attributed to land. In fact this is one way of looking at the economic productivity of a factor. Generalising this procedure it may be shown, as Clark,[2] Wicksteed,[3] Wicksell[4] and others have shown, that the economic productivity of any factor can be estimated by deducting from the total product the sum of the prices (determined according to marginal productivity) of other factors that co-

[1] *Principles* (ed. Gonner), p. 314; see also pp. 48-49; and *Ricardo's Essays* (ed. Gonner), p. 228.

[2] *Quarterly Journal of Economics*, April 1891; also *The Distribution of Wealth*, Chap. XII.

[3] *Co-ordination of the Laws of Distribution*; also, *Commonsense of Political Economy* (ed. Robbins) Vol. II, Chap. VI.

[4] *Lectures on Political Economy* (ed. Robbins), Vol. I, pp. 106-144.

operate with it. If the land owner is the hiring factor and the dosing is done in terms of labour-and-capital, then the economic productivity of land appears as a surplus. If the capitalist is the hiring factor and the dosing is done in terms of land-and-labour the economic productivity of capital appears as a surplus. The intensive working of capital or of labour as of land gives rise to diminishing returns, so that there is a no-interest margin for capital and a no-wage margin for labour, just as there is a no-rent margin for land. And since under perfect competition any of the factors may play the 'entrepreneur' there is no contradiction between the two hypotheses. In equilibrium the hiring factor secures an income which is equal to this economic productivity, that is to say, the surplus that is left over at the margin. He will not stop investment until this margin is reached, nor will he push his investment beyond it. At the margin his income will be maximised. But this income, when perfect competition exists, cannot be greater or less than what he would earn had his own resources been employed as a hired factor and paid according to marginal productivity. The entire theory of equilibrium of factor income can thus be described in terms of surplus.

But although equilibrium may be so described, the construction does not bring into full view the complex of forces operating in the economy.

Firstly, it is not an independent theory of distribution. With reference to any particular factor it tells us that when other factors are paid off it receives the residue that is left over. If there are only two factors and the value of one is determined independently, then, of course, it is unambiguous to say that the other receives the surplus. Ricardo did really proceed upon this assumption. In the simple scheme with which he started, land and labour were the only factors of production, of which labour had an absolute

supply price based on standard of living. Wages therefore were independently determined and rent was described as a surplus. This is why he was so much puzzled over the time element as a third factor.

Secondly, it does not provide a clue to the unit demand rate for a factor. According to the surplus principle, we discover the aggregate income of a group first and then the fractional income by averaging, whereas —in so far as our factor is divisible, and surely land, of all things, is divisible —it is convenient to proceed from the unit demand rate to the aggregate income, for we can then profitably watch the operation of the principle of substitution at the margin.

Thirdly, the conception of factor income as a surplus has a peculiar suggestiveness that it is no part of cost. It does not bring out the functional relationship between factor prices and prices of goods of the first order.

COST OF PRODUCTION AND RENT FURTHER CONSIDERED

1. In the last chapter we saw that the conception of land rent as a transfer income emerged from two distinct hypotheses. It was, on the one hand, a logical outcome of Adam Smith's philosophical account of value which gave rise to a real cost doctrine,—a doctrine that laid emphasis on the psychological resistances on the part of individual factors used in the process of production as the type of social cost. On the other hand, it was deduced from the assumption of a constant supply of land, — land being defined as the 'original and indestructible powers of the soil'.

In the present chapter we shall examine some further implications of these two hypotheses.

2. If by cost of production is meant the psychological disutility involved in the production of goods,—the disutility relating to individual factors taking part in production, then surely an obvious distinction appears as between 'human' and 'non-human' factors of production,—the factors that are the creation of human exertion and those that are offered by nature. The latter do not, whereas the former do, involve cost. And if the 'natural' agents are scarce and command a price in the market then their incomes present themselves as a sort of surplus. In the case of human factors, their incomes are the reflection of some mental sacrifice. In so far as the labourer helps production by undergoing physical and mental exertion, he

gets his wages. In so far as the capitalist does so by sacrificing present income for future and in some cases by undertaking a certain degree of risk, he gets his interest or profits. The payments that are made for these factors are payments against cost of production. They are payments which are necessary to lure these factors into the job, — to cover up the sacrifices which they undergo. If the condition of the market is such that the price of goods that they produce just covers these payments then price is equal to cost of production. The use of natural agents does not involve any such cost. No efforts or sacrifices inhere in such agents. Any income that accrues to them is a sort of conjuncture-income. It is a surplus above cost of production and arises out of the fact that in so far as these factors are scarce, they work in the same way as any human factor does towards limiting the supply of goods for which they are used.

It is in this sense that Senior uses the term cost of production. According to Senior, cost of production means "the sum of labour and abstinence necessary to production." Wages and profits are the price of labour and abstinence. They are a price to be paid in order to induce people to undergo these sacrifices. Rent, on the other hand, is "the surplus produce arising from the use of an appropriated natural agent." Not that rent does not 'enter into' price. But it is no element of cost of production. Senior is one of the earliest economists to emphasize in most clear terms the role of scarcity in economic valuation, and as such indeed he can claim to be one of the precursors of the modern subjective theory of value. Senior makes it clear that the value of goods arises out of their scarcity in relation to demand, and that any force that contributes to the scarcity of goods is to be reckoned with as regulating their value. A commodity can have value in the market even if no labour

is spent on it, if only it is scarce in relation to demand. When, however, it is reproducible, its scarcity is due to the scarcity of the factors that are used in its reproduction. The price of such a commodity is therefore equal to the sum of the prices of factors that are required to produce it. These factors again command a price in the market because they are scarce in relation to the demand which is derived from the demand for goods. Thus an all-embracing principle of scarcity explains the pricing of goods and the pricing of productive factors. In all this Senior's teaching ranks as an anticipation of Walras and Cassel.

How, then, does he reconcile this theory of value with his real cost concept?

The value of goods is regulated as much by labour and abstinence as by scarce natural agents that enter into their production. But whereas the limitation of supply of the former arises from the fact that they involve effort and sacrifice, the limitation of supply of the latter is due to natural circumstances. When appropriated, these natural agents would be withheld from use unless a reward were made. The necessity of using such agents in production therefore creates a divergence between value and cost of production. "The assistance of an *Appropriated* Natural Agent may render possible the production of a commodity more valuable than the result of equal labour and abstinence without such assistance. Such a commodity sells for a price exceeding the sum of the wages and profits which are sufficient to repay the capitalist and the labourers who have been employed on it. The surplus is taken by the proprietor of the natural agent, and is his reward, not for having laboured or abstained, but simply for not having withheld what he was able to withhold; for having permitted the gifts of nature to be accepted."[1] The income from such natural

[1] *Political Economy* (5th edition), p. 90.

agents is rent and is the result of the circumstances of the market creating a disproportionality between value of goods and their cost of production. And such circumstances, according to Senior, are the result of monopoly. Senior defines perfect competition as a condition in which all natural agents are universally accessible because unlimited in supply, so that no payments need be made for their exploitation. The scarcity of any such agent is an obstacle to perfect competition,—an obstacle which makes the value of goods exceed the cost of labour and abstinence and is a condition making for the emergence of rent. Rent is a generic income and relates not only to the external objects of nature, e.g. land, but also to all those productive virtues which inhere in and are appropriated by individuals, the possession of which had not called forth any effort or sacrifice on their part. All incomes from land, from inherited wealth, from superior natural talent and skill are a monopoly income and a surplus above cost of production.

Now, before we go more closely into an examination of the implication of the real cost doctrine here set forth we ought to point out that this analysis of Senior provides a type of confusion which was noticed earlier,—a confusion between scarcity and monopoly which has since crept into the writings of many economists.[1] It assumes that labour and abstinence are homogeneous elements and are delivered to the community at constant cost,—the only cause of diminishing returns being the monopoly of natural agents. In fact, however, the phenomenon of diminishing returns relates to the proportion in which factors are combined in production. It is now common knowledge that if production involves just the use of more than one scarce factor, diminishing return is an inevitable pheno-

[1] Cf. particularly Maurice Dobb, *Capitalist Enterprise and Economic Progress*, which is a recent instance of such confusion.

menon,[1] and that it is the operation of diminishing returns that creates the problem of choice in the method of production. Indeed, even if all natural agents were freely accessible,and labour and abstinence were the only two scarce factors, there would still remain the problem of choosing the proportion in which to combine them, and that problem would arise precisely on account of the operation of diminishing returns from these individual factors. To relate diminishing returns to monopoly is to use the term monopoly as co-extensive with scarcity, or —for the matter of that—as co-extensive with economic analysis in general.

There are thus difficulties on analytical grounds of drawing a distinction between the scarcity of natural agents and the scarcity imposed by psychological resistances on the part of 'human' factors of production; because both present the same case so far as the problem of pricing is concerned. What meaning, for example, can there be in isolating the extra income secured by a labourer having superior natural skill and calling it a surplus? It was only a clever expedient by which Senior tried to get out of the difficulty of explaining the differences in remuneration to different qualities of labour.

Yet this real cost or absolute sacrifice in terms of labour and abstinence has often been considered to be important as indicating the real burden thrown on the members of a community as a whole for achieving a given material benefit. Money cost in terms of reward of factors, it is held, conceals instead of bringing out their real sacrifice.

Cairnes in particular emphasizes that cost of production, in order to have any significance for economic speculation, must be understood to mean sacrifice and not reward. Not

[1] This point has been recently emphasized by Mrs. Robinson (cf. *The Economics of Imperfect Competition*, p. 330), but is already implied in the writings of Wicksteed and Knight. See Wicksteed, *Co-ordination*, Sec. 5; Knight, *Risk Uncertainty and Profit*, Chap. IV.

only is cost not identical with reward, it is the antithesis of reward. Cost means sacrifice incurred by man in productive industry, while reward is the return drawn from nature against that sacrifice, and the two things are fundamentally opposed. Economics, Cairnes urges, loses all human interest if this antithesis is not sharply drawn; and for this purpose he points to the contrast between 'social value' and social aggregate cost in terms of absolute sacrifice, and takes the proportion between the two as a mark of economic progress from time to time.

There is also in the background of Marshall's theory of costs an attitude like this; and this partly explains his sympathy for the theory that rent does not enter into cost of production.[1] Although for purposes of economic analysis he concentrates on money cost or 'expenses of production', yet there is that 'metaphysical odor' in his system, and the real cost concept peeps out here and there in it. Marshall is conscious that real costs refer to forces that are not susceptible of quantitative measurement, yet he emphasizes that they are qualitatively important in our enquiry *from the social point of view* as to "whether the cost of obtaining a given result is increasing or diminishing with changing economic conditions."[2]

[1] "....it is wicked to say" he writes to Edgeworth, "that 'Rent *does* enter into cost of production', because that is *sure* to be applied in such a way as to lead to the denial of subtle truths...." *Memorials of Alfred Marshall* (ed. Pigou), p. 435.

[2] Here is a passage which characterizes Marshall's attitude towards this aspect of social cost:

"..in the discussion of social problems, it is often necessary to inquire whether certain businesses, which may or may not be adequately remunerative to the persons concerned, are worth what they cost to the country or the world: and in this connection the term cost of production refers to real cost. The *real cost of production* of a thing is the aggregate of efforts and sacrifices which are incurred in its production. Thus the work of very young children in factories, even though paid for in money at the full market rate, is seldom worth its real cost: the satisfactions, which are derived from its

Viewed in this light cost refers to the sacrifice on the part of the individual factors taking part in production. It raises the question, how much it costs to *the labourer* or to *the saver* in furnishing productive services adequate for a given amount of output. These costs by themselves are purely subjective and depend upon individual sensibilities. From the point of view of a labourer it is a question of his individual scale of preference as between real income and leisure, or as between different kinds of occupation. From the point of view of a saver it is a question of his scale of preference as between present income and future income. Sacrifices of this kind vary from person to person and with regard to the same person they vary at different levels of income. The so-called 'disutility' is not measurable or comparable as between different persons.[1]

Nor are such estimates helpful or necessary for a study of the objective relationship between goods and the role of cost in that context. What is relevant there is not the cost to the individual resource owner but cost to the society in the sense of displaced alternative. What it costs to the society in employing a particular amount of productive resources for one good is the value of other goods which these resources might produce. Cost of one thing is the value of the alternative surrendered. It is not the *measure* of anything which is absolute ; on the contrary, it implies a comparison of what *is* with what *might be*, and thus expresses the actual in terms of the potential. The relation

contributions to production, are not worth the social cost of child life spent in grievous and depressing toil, and without any adequate education to prepare for the duties of after life." (*Industry and Trade* Bk. II, Chap. I, Sec. 2, p. 183.).

[1] Some discussion on these psychological preferences and the kind of surplus that may arise out of an 'exchange' of this sort will be found in the subsequent chapters.

between cost and value lies in so far as the value ratio between goods corresponds to this displacement cost ratio.[1]

Thus the cost of labour in any particular employment is not the absolute measure of the labourers' 'toil and trouble', but *is equal to* the value that is surrendered on account of its being tied up in that particular employment at the expense of other alternative employments. Similarly, land in so far as it is capable of being used for alternative purposes involves cost in any particular use, and this cost is equal to the amount of service which it might have rendered in other alternative directions.

This at once explains why factors are remunerated, and also why different kinds of labour and land are differently remunerated. Factors are remunerated because of their scarcity in relation to demand,—because if they do not get adequate remuneration in one employment they would transfer their services to other employments. Production has been compared by Böhm-Bawerk to a 'giant pump'. "Every branch of want has its separate pipe sunk down to the great reservoir of productive powers, and competes with all other branches of want in trying to draw its supply by suction from that reservoir."[2] But as the reservoir contains limited quantities of productive powers, their employment in

[1] Displacement cost, it should be noticed, is expressed in value terms and not in terms of product. Although it is somewhat of a tautology to say that the displacement cost ratio in value terms conforms to the value ratio (all equations are a tautology !), yet, in order that the principle may be universally applicable it should better be so stated. See, e.g., Robbins, *Some Aspects of the Theory of Costs*, Economic Journal, September 1934, where it is shown, among other things, that often when the commodities in question are not produced with the same technical resources, or, when the proportion in which factors are combined in production (technical co efficient) is fixed but different for each of them, the technical displacement cost principle does not explain value.

[2] The *Positive Theory of Capital* (English Translation by William Smart), p. 230.

one branch of want means that the other branches are deprived. In the struggle among competing branches the fittest survive, and the productive powers are carried off by those that have stronger powers of suction. This increases the supply and lowers the relative marginal utility of the stronger alternatives, and reduces the supply and raises the relative marginal utility of the others until a stage is reached at which they coincide with one another. As a result of this process of adjustment the productive powers get a value imputed to them in such manner that the cost ratio in terms of their reward may correspond to the value ratio between goods. The psychological preferences of individual resource owners are significant in this context only in so far as they form part of our data,—in so far, that is to say, as they determine what amount of a particular factor should be available for employment.

This cost principle may be derived from—indeed it is implied in—the studies of the classical economists from Adam Smith down to Senior in so far as these studies fix attention on price as a ratio; although, as is evident from what has been said above, even there it is conceived in terms of simple production. In Ricardo it is homogeneous labour, in Senior it is a combination of labour and abstinence, that forms the standard for cost comparison. But in fact the principle applies to all factors that are scarce and that have got alternative uses.[1]

[1] A precise formulation of the principle of cost as displaced alternative is due to Wieser. See his *Natural Value*, Bk.V. See also in this connection, Wicksteed, *Commonsense of Political Economy* (ed. Robbins) Vol. I, Chap. IX; Knight, *Fallacies in the Interpretation of Social Cost*, Quarterly Journal of Economics, 1924; *Suggestion for Simplifying the Statement of the General Theory of Price*, Journal of Political Economy, 1928; Haberler, *The Theory of International Trade*, Chap. XII. Robbins, *op. cit.*, pp. 2-6.

Recently Prof. Knight, following Viner who has still a leaning towards the real cost concept (cf. the latter's article on *The Theory of Compa-*

The twin principles of Scarcity and Substitution relevant to the problem of cost and factor prices apply to land as much as to other factors. In the first place, it is to be remembered that money payments that are made to factors in a particular industry are not the *measure* of anything absolute. They are payments that are calculated to equalise the marginal value returns from a unit of resources in different industries; or, in other words, to equate the marginal cost ratio with the value ratio between goods. In this respect land and other natural agents stand on the same footing as labour and capital. Again, to an entrepreneur (or to society in relation to which the entrepreneur in a system of free enterprise acts as a sort of middleman) seeking to minimise cost through substitution of one factor for another as their relative prices in the market warrant, land occupies the same place as any other factor.

3. Does any new principle emerge on account of any peculiarity in the condition of the supply of land? The Ricardian hypothesis, as we have noted, is that the supply of land is absolutely fixed and invariable. It just follows from the way he defines land. Has this assumption an economic

rative Costs in *Welt-wirtschaftsliches Archiv.* Bd. XXXVI, 388-414), has pointed to a limitation of the displacement cost principle: (see his review of *Commonsense of Political Economy* in *Journal of political Economy*, 1934). Wieser's principle of cost provides an explanation of relative values if (*i*) individual resource owners are indifferent as between the relevant occupations, or if (*ii*) even though the occupations are not equally preferred, the resource owners make up the difference by 'voluntarily' working more in the preferred occupation and thus equalising their money incomes; but if the 'advantages' of occupation are unequal, and the difference is not so made up, then, it is suggested, the principle does not hold.

As, however, Prof. Robbins points out, the difficulty is solved "if the other advantages and disadvantages are treated as joint products." *Op. cit.*, footnote on p. 3. See also in this connection the important article of P. N. Rosenstein-Rodan, The Role of Time in Economic Theory, *Economica*, Feb. 1934, p. 87. Cf. Appendix, below.

significance warranting a distinction of land rent as a separate category from other forms of income?

From what has been said above it is quite clear that our analysis of the functional relation between factor incomes and the relative prices of goods is independent of any assumption concerning the variability of the 'total' supply of factors. In a sense all factor incomes are derived from the value of goods. The Imputation theory of Distribution points to this. On the other hand, if, in any given situation, the price of any good exceeds or falls below the expenses of production, it only indicates a mal-adjustment in the distribution of factors as between different productive employents; and, given time, the process works out in such manner that prices correspond to costs. In this sense, cost, which is made up of payments to factors of production, does influence the price of goods.

Now, in providing a description of equilibrium in the distribution of factors it is assumed as a first approximation that all factors of production are fixed in supply. And it is certainly true that they are fixed at any given moment of time which is taken as the starting point. Further, one of the conditions necessary for the maintenance of the equilibrium that is envisaged is that the flow of factor supplies continues to be constant over time. Any change occurring in it is a change in datum which is calculated to bring about a new position of equilibrium. Within these conditions the amount of any factor available for a particular line of production is determined by the resistances offered by competing lines. In all this there is no warrant for putting land as a special category.

But what in the ultimate analysis is it that determines the 'total' supply of a factor? Is not there a price which is necessary, absolutely speaking, for maintaining a stationary supply of certain productive factors,—such that when

that price is varied there follows a corresponding variation in the supply of the relevant factor?

It is in connection with this ultimate determinant of the total supply of productive factors that the Ricardian theory of rent as a transfer income is considered to have special significance. Labour and capital are variable and have got a maintenance price. The supply of land is rigidly fixed and is independent of any income attaching to it. Wages and interest are items in the 'social' cost of maintaining intact the supply of labour and capital. Rent is a surplus depending upon the demand for land in relation to the given stock.

Marshall lays particular emphasis on this aspect of the classification of factor incomes. He recognises that it is misleading to say that rent does not enter into cost of production. From the point of view of an individual farmer deliberating whether he should apply a little more of land or of other factors to his organisation, rent constitutes an element of cost as much as other payments.[1] With reference to a single crop again, land, in so far as it has a transfer price, behaves exactly like other factors; for, the rent which a field is capable of yielding in alternative employments constitutes, even as wages and interest do, a part of the expenses of the crop for which it is actually used. (The only proviso that Marshall adds in this connection,—no doubt an important proviso,—is that there is not necessarily a 'numerical relation' between the transfer price and the actual earnings of land). But—and this is what Marshall's theory of rent ultimately turns on—from the point of view of society land is a fixed quantity and has no supply price. The services offered by land have a constant flow over time irrespective of what is paid for them, while so far as the variable factors are concerned, the rate at which they are

[1] "Land is but one form of capital to the individual producer". *Principles*, p. 430.

supplied would depend upon the rate at which they are remunerated.

Marshall's difference with Clark and Cassel on the question of the distinction between land and capital ultimately turns on this last condition of what determines the equilibrium supply of factors of production from the social point of view. With the latter all factors of production have by hypothesis a constant flow.[1] With Marshall land is by definition constant, whereas other factors are variable and have a schedule of supply prices. In his analysis of 'statical equilibrium' Marshall fixes his eyes on some central point where the economic forces come to rest; but the variability of those forces is recognised. The forces themselves tend to change, but equilibrium is reached as a result of a balance between these forces.[2] This does not rule out the possibility of a variation of labour and capital. But land being nature's gift is a constant quantity. Under such conditions the income from labour and capital reacts on their supply, while the income from land has no reaction on its supply. It is a surplus. A tax on this surplus merely brings about a redistribution of wealth, since it leaves the supply of land and hence the flow of National Dividend intact. But the supply of labour and capital being poten-

[1] See Clark, *The Distribution of Wealth*; Cassel, *The Theory of Social Economy*, p. 34.

[2] Cf. Marshall's letters to Clark, *Memorials of Alfred Marshall* (ed. Pigou), pp. 413-415; *Principles* (8th ed.) p. 369.

See in this connection a very interesting discussion on 'A Certain Ambiguity in the Conception of Stationary Equilibrium' by Prof. L. Robbins, Economic Journal, June, 1930. In the course of this paper Prof. Robbins points out the two different interpretations put upon stationary equilibrium by Clark on the one hand and Marshall on the other. Marshall's conception is one "in which the condition of stationariness is the resultant of the balancing of forces tending to change," and the Clarkian conception is one "in which the factors of production are stationary by hypothesis", equilibrium being attained within these conditions. See also F H. Knight: *Risk, Uncertainty and Profit*, p. 125n, 142-43n.

tially elastic at the point of equilibrium, a tax on wages and/or interest leads to a reduction in the supply of labour and/or capital, and thus produces an adverse effect upon the National Dividend.

Further, as the adaptability of factors to changing circumstances may be slow or rapid, the variable factors also admit of subdivision. If over a particular period one species of such factors has an inelastic supply, its income shows characteristics of rent and is a surplus, but it ceases to be so if, when the period is extended, its supply becomes elastic. Again, over that extended period the supply of another species may be fixed and its income may appear to be a surplus, but in a like manner it ceases to be a surplus if, when the period is further extended, its supply becomes elastic. These incomes are called Quasi-rent because they are not pure rent (surplus) and yet show the characteristics of rent in relation to a period over which the supply of the relevant factors is inelastic. A recognition of the distinction between these quasi-rents and rent proper, Marshall urges, is important both scientifically and in relation to the practical wellbeing of a society. Scientifically it is important to abstract from the composite income that which is 'true rent', that is to say, that element of it which is attributable to the 'original and indestructible' productive powers.[1] And in relation to practice it may guide the policy of the public authority. Once the institution of private property has been allowed to stand, the appropriation of both rents and quasi-rents would produce similar effects in destroying security and retarding economic progress. But if from the beginning the state were to retain to itself true rents, then no such disturbance would result.[2]

[1] "Economics learns from physics to reason about pure elements, though they are rarely isolated in nature". *Principles* (8th ed.), p. 421.

[2] *Ibid.*, pp 802-03.

Now, there are no doubt certain materials of production which are provided by nature and available for use without human industry. Certain properties of land are 'originally' there. But whether they are 'indestructible' is surely questionable. If they are not, then, in an old country at any rate, maintaining land intact presents a problem no less urgent than maintaining capital intact. This, however, is a question of brute fact. Analytically there cannot be any quarrel with Marshall's assumption as such; and indeed who realised more than Marshall did, with what degree of caution the assumption should be taken?

What, however, is important to notice is that the notion of an absolute supply price is not fundamental in economic analysis, nor is it possible to say with any degree of precision from *a priori* argument to what extent and in what direction even labour and capital should respond to variations in their prices.[1] In an exchange economy where particular attention is given to the employment of resources for pecuniary uses, — variation in the 'total' supply of factors is to be observed in relation to their potential uses

[1] Cf. D. H. Robertson, Economic Incentive, *Economica*, Oct. 1921, reprinted in *Economic Fragments*.

Following classical lines Marshall gave a concept of an absolute supply price of labour and capital with reference to 'the cost of rearing and training' men from whom these factors come. See *Principles* (8th ed.), pp. 576-77. A more explicit reference to an absolute supply price of capital and labour based on different standards of living in different classes which had appeared in an earlier edition (2nd ed., bk. VI, pp. 557-58) was deleted in later editions. Now, of course, the standard of living in certain cases may be rigid enough to exercise some influence on the supply of labour and capital. But it is difficult to see how it works. If, for example, a particular set of people is used to a certain standard and would cling to it, then a lowering of income per unit of labour may be followed by a restriction of their numbers (reduction of supply) or a tendency to work more (increase of supply). In fact the standard of living in general is not as rigid as that. It is essentially an *effect* of the size of population in relation to the degree of economic co-operation.

for non- pecuniary purposes. And there the difference between land, however defined, and other factors is at best a matter of degree. The equation between the marginal utility of leisure and real income decides the equilibrium supply of labour. The equation between the marginal utility of present income and future income decides the equilibrium supply of capital. Similarly the equation between the marginal utility of pecuniary product and non-pecuniary product decides the equilibrium supply of land.

4. The foregoing discussion will have made it sufficiently clear that the supply price of a factor of production is to be reckoned with reference to an industry and not to society as a whole, and that this supply price consists in the value return which it is capable of yielding in alternative employments. The supply price of a factor used in the production of x is the value of y which it might produce when x and y are open alternatives for which it might be applied indifferently. And, it has also been seen that when a factor is capable of such alternative uses, that is to say, when demand for it is 'composite', its income is a part of the cost of that good for which it is actually used, and further that payments to such factors are necessary to equalise the marginal cost ratio with the value ratio of goods.

If, however, a factor has a single pecuniary use, then a new principle emerges. Although, if it is scarce, a payment to such a factor is necessary for the equalisation of the marginal costs of different entrepreneurs employing it, yet such payments are not 'causal' in relation to the relative prices of goods. From the point of view of the industry to which it is attached, the factor has no supply price, and the income which it secures on account of scarcity depends upon the condition of demand for the good which it produces. This income is essentially a price-determined surplus. Following Wieser we may call such factors 'specific produc-

tive means' as distinct from 'cost productive means'. Ricardo's proposition that rent does not enter into price does hold good when land admits of a single use,—in other words, when it is 'specific', although it loses its applicability when the uses of land multiply. It is not a rigidity in the 'total' supply of a factor but rather a rigidity in its supply from the point of view of a particular industry arising out of specificity that is responsible for the emergence of a price-determined surplus of the sort contemplated by Ricardo.

The existence of specific factors brings in certain rigidities in the economic system and introduces special problems for economic analysis. In our study of the effect of 'variations' upon the supply condition of goods and the distribution of income, for example, we have to take account of whether the factors employed for productive purposes are specific or non-specific. Variations in the relative demand for goods arising from changes in taste, taxes or bounties, opening up or closing of trade connection with foreign countries, etc., affect specific factors in a different way from non-specific factors. If on account of any of these causes there is a shrinkage of demand in one industry, it is, of course, accompanied by an expansion of demand in some other industry or group of industries. In such a case the non-specific factors, when thrown out of the depressed industry, are absorbed in others which are stimulated. But so far as the specific factors are concerned, such transference is not possible, and those of them that are specific to the depressed industry earn less than before, whereas those that are specific to the stimulated lines earn more than before.

Here again the difference between land and other factors is more a difference of degree than of kind. Specificity is not the peculiar property of land, but may attach more or less to any factor of production.

In fact, specificity is itself a matter of degree. It is highly improbable that any factor should be absolutely specific. The degree of specificity is indicated by the excess of the actual earning of a factor over that which it might earn in the next best possible use. And the surplus that is considered to arise out of specificity relates not to the total earning of the factor in question but only to that part of it which is beyond its transfer price.

Further, it should be noted that although to the extent that any unit of a factor group is specific it earns a surplus, yet taking a factor group as a whole its price may influence the price of its product if only *some* of the units are capable of alternative uses. This consideration is important in as much as the units of a factor group although homogeneous from the point of view of the industry in which they are actually employed, may yet be heterogeneous from the point of view of other industries. This is often true of labour and may be true of land, too. As regards capital, so long as it is 'free', it is not specific at all ; it may only become specific—and specific to a high degree—when 'fixed' in machines of different kinds, in which case, however, each kind would contain units which are more likely to be homogeneous in every respect. But specificity of this sort is a question of time.

This leads us to an important consideration,—that of the role of time in the problem of specificity,—a consideration from the point of view of which Ricardian land (original and indestructible productive powers) is at last seen to have a theoretical if not a practical importance.

Once capital is 'fixed'—whether in human agents or material equipments—and adapted to a particular line, it cannot be transferred to other lines. On the other hand, such capital equipments require time for construction. Hence in the short period the alternatives are 'closed' to such

capital, so that it does not come under the opportunity cost principle.[1] Its income in the short period is regulated solely by demand and is a surplus. Marshall's analysis of supplementary cost and quasi-rent *in relation to an individual industry* points to this phenomenon. When, however, the period is long enough to make full adaptation possible, capital becomes 'free' and transferable and hence variable in relation to particular lines of employment. Its income then has no longer that character of a surplus. From the long run point of view, supplementary costs take the character of prime cost and quasi-rents 'shade into' interest.

At first sight it appears as though the length of time required for such adaptation should be so great as to cause a damaging degree of rigidity in the economic system,—as though the time required for adaptation to an upward turn would be the period of construction, and the time required for adaptation to a downward turn would be the period taken for fixed capital to wear out. If we had made a rigid division of time and circumscribed it by boundary lines that would seem to be the case. But time is a continuous flow having, so to say, no beginning and no end, and adaptation is a continuous process. What is relevant there is a *rate* of demand and a *rate* of supply of productive services. Moment by moment[2] investors are making anticipations of demand and trying to adapt their plans accordingly; and as there are always supplies of capital and labour floating in the system and ready for investment, adaptation becomes quicker than appears at first sight.

If, however, one of our specific factors is 'original' and 'indestructible',—if, that is to say, it is not capable of renewal and does not require replacement, then it would maintain

[1] Wicksteed, *op. cit.* Chap. IX.

[2] By 'moment', of course, is meant a period just sufficient to enable entrepreneurs to revise their plans.

a constant flow through time and its income would strictly be a surplus.

This, of course, is more a theoretical possibility. 'Original' factors of production,—natural land and raw labour,—are seldom specific to any significant degree. What renders such factors specific is the investment of capital turning their properties into distinct channels, and in so far as this is the case, possibility of transference of land into alternative lines is a question of time.

This question of time and adaptation of the economic process takes us far into the field of economic dynamics, certain observations on which will be made towards the end of this Essay.

APPENDIX

ON THE ULTIMATE NATURE OF COSTS

What is the ultimate explanation of the money costs or 'expenses' of production of goods to which reference has been made in the previous chapter? Does the ratio of money costs correspond to the ratio of 'dis-utilities' involved in the production of goods—or does it correspond to the ratio of 'opportunities' surrendered? According to one view, the relative money cost of production of goods is the reflection of the relative quantities of 'pain' or 'disutility' involved in their production. According to the other, it is the reflection of the relative sacrifices involved —in terms of goods or values thereof—in view of the employment of a given quantity of resources. Suppose that the cost in terms of money of producing one unit of x is twice that of producing one unit of y, so that the ratio of money cost of x and y is 2 : 1. Are we to say, then, that the ratio of real cost in terms of 'pain' of the two goods is also 2 : 1,—or merely that the production of one unit of x requires some resources the employment of which involves a sacrifice of 2 units of y,—in other words,—that the alternatives obtainable from a given amount of resources are one unit of x and two units of y? The former is the so-called 'real cost' approach and the latter is the 'opportunity cost' approach.

Now, it may be asked,—why seek for a non-monetary explanation of costs at all? The real world being a world of prices and money costs, why not be content with just an 'empirical' account denoting the relation between these two? It is indeed true that the producer takes factor prices in the market as a fact and endeavours to use such methods as reduce the money cost of an article to a minimum. There is also a tendency under competitive conditions of prices of goods to

conform to these money costs. And of course all these form interesting studies. Yet they do not tell us enough. An analysis of the ultimate nature of costs contains more profound principles. The significant point in this analysis is that it seeks to indicate the condition of maximum economy in the real sense.

According to the real cost doctrine minimum money cost of production implies minimum real cost; and if prices conform to money costs, it follows that goods are produced at the minimum sacrifice in terms of pain. Perfect competition, therefore, ensures maximum economy. If, on the other hand, the money cost ratio fails to represent the real cost ratio, there arises a need for interference. If the price as well as the money cost ratio of x and y is 3 : 1, whereas the real cost ratio is 2 : 1, the process of increasing the output of x and reducing the output of y so as to make the price ratio conform to 2 : 1, is a process that results in an economy of effort. The equilibrium price ratio, 3 : 1, and the output that goes with it are thus uneconomical.

The same fundamental theorem emerges also from the opportunity cost principle, although the elements considered there are more objective in character. Given the quantity of resources and given the alternative possibilities of the employment of those resources, maximum economy requires that the price ratio between goods should be the same as their opportunity cost ratio. If, as in the above example, the opportunity cost ratio of x and y is 2 : 1, and the price ratio is 3 : 1, increasing economy is effected as resources are transferred from y to x until the price ratio is brought down to 2 : 1. Now, under competitive conditions the price ratio tends to be equal to the money cost ratio. If, therefore, conditions are such that the money cost ratio just reflects the opportunity cost ratio, freedom of enterprise leads to maximum economy—for, in such circumstances, the correspondence of the price ratio to the money cost ratio does imply its correspondence to the opportunity cost ratio. Taking the simple case of constant returns, if, as indicated in Figure 1 below, CC′ is the Production Opportunity curve and tangent of the angle OC′C represents the opportunity cost ratio, showing the alternatives in terms of x and y that

can be had with a given amount of resources, the total product stands at a maximum just when this ratio coincides with the consumers' preference ratio or the market rate of exchange.

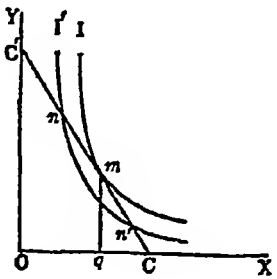

FIG. 1.

If the Consumption Indifference curve, 1, touches CC' at m, then m is the point of maximum economy,—or in other words, the given quantity of resources yields maximum output when Oq of x and mq of y are produced. Any other output represented by the co-ordinates of, say, n or n' would be less preferred. If, therefore, the money cost ratio to which the price ratio corresponds fails to be equal to the opportunity cost ratio, the resulting output is less than maximum.

Now, if there is only one scarce factor entering into production,—homogeneous labour—either of the above principles may serve our purpose. An hour's work may be taken as involving a definite quantity of pain, and the pain cost of one commodity can very well be compared with the pain cost of another, and this ratio turns out to be just the same as the opportunity cost ratio :

2 hours' work produces one unit of x
1 hour's work produces one unit of y

The production of one unit of x involves double the pain involved in the production of one unit of y. The pain cost ratio is 2 : 1.

Again :

Given 2 hours' work, the alternatives open are one unit of x and 2 units of y. The production of one unit of x involves a sacrifice of 2 units of y. The opportunity cost ratio is 2 : 1.

This holds good of Adam Smith's imaginary primitive society. Complications arise when production is complex,—when, that is to say, a number of scarce factors come into the picture; for, in that case, the pain involved in different operations cannot be estimated. The psychological discomfort, if any, involved in "abstinence" cannot be compared with that involved in "work"; nor can different kinds of work performed by different labourers be brought under a common standard in terms of pain.

Of course, even in the case of complex production, no damage is done to the real cost principle, so long as the proportion in which different factors are employed in different industries is the same. For, here again any of the factors can be taken to indicate the cost ratio, and indeed as estimated in terms of an hour's work the pain cost ratio would be the same as the opportunity cost ratio.

2 hours' work and 4 units of capital yield one unit of x.

1 hour's work and 2 units of capital yield one unit of y.

The cost ratio in terms of any of the factors is 2 : 1, and there is no cleavage between the real cost principle and the opportunity cost principle.

The real cleavage between the two principles arises when the proportion is not the same. When factors are employed in unequal proportions in different industries there is no means of discovering the real cost ratio. You cannot compare the real cost of, say, 2 units of labour *plus* 5 units of capital with the real cost of one unit of labour *plus* 2 units of capital except by attaching arbitrary indices of disutility to the factors.

On the other hand, an hour's work and a unit of capital by themselves are perfectly definite concepts and can be related to the products that they are capable of yielding. Starting off with a bundle of resources—whatever its contents might be[1]—

[1] The contents, of course, are not to be arbitrarily chosen. They would depend upon the relative scarcity of factors, or, in other words, upon the relative marginal productivity.

one can discover the possible alternatives open and can deduce the opportunity cost ratio without any reference to the relative disutility of different factors. The principle is perfectly general and does explain conditions of simple production as well as complex production of any form. Unequal proportions in the use of factors in different industries merely indicate that some factors

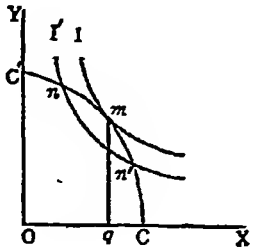

FIG. 2.

are 'specific'. Specificity leads to diminishing returns as the product of one good is increased at the expense of the other. This, however, does not cause any trouble. For, it only means that the marginal opportunity cost ratio varies as against that good whose output is increased, and the Production Opportunity curve is concave to the axes. The condition of maximum economy is again equality between the price ratio and the marginal cost ratio. The point, m, in Fig. 2 above, where the Consumption Indifference curve, 1, touches the Production Opportunity curve CC' (in this case concave), is the most preferred point, indicating that the combination, Oq of x and mq of y, is the maximum product yielded by the given bundle of resources. The concavity of the Production Opportunity curve signifies diminishing return which,—translated in terms of the opportunity cost principle—means that as one moves on along the curve producing more of one commodity, the rate of substitution for that commodity in terms of the other—as indicated by the slope of the tangent to the curve—increases. This is nothing but to say that the cost of having a little more of one commodity in terms of the other increases as resources are transferred in one direction. In Marshallian terms—this happens

because of a disproportionate use of factors. It is now easy to see how this disproportionate use is the result of the existence of some 'specific' factors. As we move along the curve in favour of one commodity, transferable resources are shifted from the other line, but they have to be worked with the given quantity of those factors which are specific to it.

What is the counterpart of all this in money terms? It can be easily seen that in the case of factors that are fully transferable one important condition that needs to be satisfied in order that the money cost ratio should correspond to the opportunity cost ratio is that each should have the same price in different industries. In the more complex case where specificity is allowed, the condition necessary for the correspondence of the two ratios is that in every industry the specific factors should be paid according to their marginal productivity,—or, one might say, according to the 'rent' that they yield. As more of x is produced, the additional cost of every increment of x gets on increasing in terms of y if there are some factors specific to x, if, that is to say, the necessary resources cannot all be diverted from the y industry. This rise in the cost of x will have its expression in absolute money terms if there is corresponding increase in the price of the factors specific to x, the transferable factors having the same price in both the industries. If, on the other hand , the 'rent' of specific factors is kept at a lower level, the money cost ratio will be more favourable to x than is warranted by the opportunity cost condition. Equilibrium will take place at a position of less than maximum product. This, by the way, is the economic justification for maintaining a market price for specific factors. Even though the remunerations that market conditions offer to specific factors appear as rent,—a surplus, from the point of view of society—in the sense that the employment of the services of their receivers does not involve any sacrifice in terms of an alternative product,—the producers as such must be made to regard them as part of their expenses in order that the actual expense ratio may not diverge from the opportunity cost ratio.

All these conditions, as is well known, are realised under perfect competition. The competition among producers brings it about that prices are proportional to money costs,—prices of factors including specific factors correspond to their marginal productivity and distribution of factors as between products is such that the money cost ratio corresponds to the opportunity cost ratio.

The opportunity cost principle is thus based on a more objective analysis and has nothing to do with the distinction between the human factors and the non-human factors of production,—a distinction which begs the most debatable question of the comparability of disutilities attaching to different factors,—of whether work gives pleasure or involves pain, and so on. The cleavage between the real cost principle and the opportunity cost principle in the theory of production is analogous to the cleavage between the older utility principle and the modern 'choice' principle in the theory of consumption and value.

Let us, however, turn to objections. An old objection which was raised by Edgeworth is that while the opportunity cost principle explains market phenomena on the assumption of a given supply of productive resources, it breaks down when the question is one of ascertaining what quantity of resources will be available to the society for productive purposes. The quantity of land is given by nature; but the determination of the available quantity of labour involves a comparison of utility of real income and the disutility of work. The answer to this objection given by Professor Robbins is decisive. Indeed the same solution can be put with greater advantage in the language of the opportunity doctrine; for the total availability of labour depends upon the labourers' choice between 'real income' and leisure. 'With greater advantage'—because one fundamental point emerges from this principle, namely, that even the total quantity of Nature's gifts, land, for example, which is available for the production of real income,—marketable goods, that is to say, cannot be said to be rigidly fixed—depending, as it does, upon the choice of the holders as between 'economic' and 'non-economic' uses.

A more subtle objection to the principle is that which has been raised recently by Professor Viner,—an objection that has led to the famous recantation of Professor Knight. It is argued that so long as the relevant occupations are equally attractive the principle holds, but that it breaks down when, as generally is the case, different occupations offer unequal attractions for a factor. For in that case the same amount of work, for example, will involve unequal 'costs' in view of differences in 'irksomeness,' even though labour is otherwise homogeneous.[1]

Now, it is true that the opportunity cost theory starts as a first approximation with the assumption that the factors are indifferent in the matter of choice of occupations. If x and y are products of industries which are equally attractive to our resource owners the maximum principle as deduced from equality between the price ratio and the opportunity cost ratio holds unambiguously. And of course it is in order to bring into relief this salient point that this first approximation is resorted to. If, however, the y industry is more attractive than the other, there arises yet another question of a preference between 'real income' and 'other advantages'; and there is no knowing how people will react to this. They may work more in the y industry and equalise money incomes, or they may accept lower money incomes and enjoy 'other advantages' attaching to the industry.

The real cost theorist would resolve this difficulty by taking an hour's work in the y industry as involving less cost than in the x industry, thus arguing that if money wages differ in the two industries that is a reflection of a difference in real costs.

This is all right so far as it goes. But at the same time the problem can be also thrown in terms of the opportunity cost principle. Whereas in the above solution an element is deducted from the cost item,—here the 'other advantages'

[1] Strictly speaking, this applies to all kinds of resources. Even property-owners do take into account differences in 'risk' attaching to different enterprises.

will have to be added to the income, being regarded as a joint product of the y industry. And if that is allowed, it can be very well seen that whatever may be the preference of the resource owners—equalisation of the price to the marginal expense ratio does ensure maximum product. When the preference scheme of factors is such that the greater attractiveness is wholly compensated by extra work, and money incomes are equalised, the maximum product runs in terms of the commodities themselves and the principle is unambiguous. Where, however, money incomes are not equalised, the product maximised is in terms of x and y *and* 'other advantages'.

Professor Viner, however, objects to this 'product' terminology which includes such things as 'advantages of occupation'. Referring to Professor Robbins' treatment of 'other advantages and disadvantages' attaching to different occupations, Viner says, "By calling the excess of pleasurableness of occupation A over occupation B a 'joint product' of A..... the product terminology is retained while proper account is taken of the significance for prices of choices between *other alternatives than products.*"[1]

What, then, one might ask, would Professor Viner mean by the term 'product'? Does it consist only of material objects? Of course, not. Of all persons Professor Viner could not be considered to be still under the physiocratic influence. If, on the other hand, it consists, as it does, of all exchangeable things, then it might surely include these 'other advantages'. If I prefer to pay a penny more for an article of the 'same kind' because my shopkeeper has pleasant manners, then this immaterial benefit does enter into the circle of exchange, and there cannot be any harm in calling it a joint product in the same way as wool and mutton are joint products, and a sheep yielding more (or a better kind) of wool sells at a higher price, although it yields the same amount of mutton as the others do.

This, however, is a matter which ought not to create much trouble. It is of little consequence whether in view of differences

[1] Viner, *Studies in the Theory of International Trade*, p. 525. My italics.

in the attractiveness of occupations the element of advantage and disadvantage is considered in relation to 'resources' or in relation to 'income'. The difference turns out to be one of form rather than of substance. The same remark applies to that version of the 'real cost' theory where by 'real' is implied resources in *kind* as distinct just from *money*. Indeed it does not matter whether the economic process is viewed as the employment of minimum resources for a given income, or as the achievement of maximum income against given resources. The so-called 'real cost' analysis lends itself more readily to the former view, and the opportunity cost analysis lends itself more readily to the latter view.

A genuine divergence between the two theories arises only when by real cost is meant psychic pain or 'disutility', capable of being measured and compared. This is the doctrine that comes from Senior, and it is against this that 'Wieser's Law' is a reaction.

GAINS FROM EXCHANGE

CONSUMER'S SURPLUS

CHAPTER VI

GAINS FROM EXCHANGE: EARLIER NOTIONS

1. The theories of surplus that we discussed in the preceding chapters, all relate to the problem of distribution. They relate to the problem whether any part of the flow of national income could be considered as a surplus above 'producer's costs'. In this and the following two chapters we shall discuss the implications of the theory of Consumer's Surplus.

Broadly speaking consumer's surplus signifies a gain from exchange. It implies that the satisfaction which a consumer derives from an economic good is greater than the sacrifice in terms of other goods which he makes in order to secure it. The consumer's (psychic) income exceeds his (psychic) cost. When, as in the case of 'buying and selling' (as distinct from barter), the consumer pays money in exchange for a good, his cost is represented by the utility of other goods which that money might have secured for him. And if the utility acquired exceeds the utility forgone, there results a gain from exchange. That which is received stands higher on the scale of preference of the consumer than that which is foregone, and his act of exchange does give rise to a gain or surplus. The ratio of prices of the goods exchanged corresponds to the ratio of their marginal utilities and not to their total utilities. Until that equality between the ratio of prices and the ratio of marginal utilities is established the consumer secures a gain,—for, that which is secured is preferred to that which is surrendered; at the margin he is indifferent, and beyond it he does not

move. In the case of 'buying and selling', if money is taken as a medium of exchange and as representing a complex of commodities, the gain from exchange is represented by the difference between the utility of that good which *is* secured for a given sum of money and the maximum utility that *might be* secured from all those goods from which that money is withdrawn. Here also the fact of the consumer's preference suggests that the utility secured is greater than the utility foregone. It is by so distributing his resources that the relative marginal utilities of goods equal their price ratios that the consumer attains his maximum position. And as different individuals have different scales of preference (in other words different comparative utilities), in any exchange transaction both the parties to it may in their individual ways derive a surplus utility.

2. It should be made clear at the very outset that the concept of consumer's surplus forms a plane of discourse entirely distinct from Ricardo's differential surplus theory of rent. There may be some superficial analogy between the two concepts. Ricardian rent is based upon the phenomenon of diminishing returns and consumer's surplus is based upon the phenomenon of diminishing utility, and both these phenomena may be shown to be manifestations of the same principle. So close does this analogy appear to be that more than one economist have argued them as aspects of the same problem. Pantaleoni particularly emphasizes this analogy and suggests that both consumer's surplus (in his own terminology 'residual utility') and Ricardian rent are special cases of a 'general law of rent'.[1] While Pantaleoni conceived both rent and 'residual utility' as a sort of

[1] Cf. *Pure Economics*, p. 156n. Cf. also Jannaccone, *Biblioteca dell' Economista*, IV[a] ser. Preface to Vol. IX—X: "Rendita del consumatore a rendita del produttore ora, invece, ci appaiono come due aspetti d'un fatto istesso". p. 64.

differential surplus and presented the two concepts as species of the same genus, Wicksteed tried to show that none of them were so. Wicksteed in fact gave a finishing touch to his Marginal Productivity theory by describing the so-called consumer's rent as the differential co-efficient of *psyche* as a factor of production of satisfaction,—precisely on a par with land rent.[1]

But the analogy is only superficial. Ricardo's 'differential surplus' is a mere transference of produce, and cannot be used as an index of an increase or decrease of prosperity or welfare of a society; whereas a change in consumer's surplus is meant to signify, other things remaining the same, a corresponding change in the 'net benefit' of a society. Whatever the character of land rent might be, it has to be sought for within the national dividend. Whether it is conceived as a differential surplus or an income determined by the marginal productivity of land, it certainly forms part of the joint product ready for distribution. Consumer's surplus, on the other hand, relates to a surplus *outside of this distributend*. It raises the question whether in equilibrium conditions, while the total product is absorbed in remunerating the different factors of production, our economic subjects yet enjoy a surplus of *satisfaction*. Ricardo's differential surplus due to the 'original and indestructible powers of the soil', if it could be abstracted from other incomes from the beginning, would appear as a peculiarly suitable object of taxation. Whether or not it is so, whether at all it is fruitful to look upon land rent as a differential surplus,—questions which occupied us in the preceding pages,—is another thing;—but that this is the significance of the historic controversy centring on the proposition—'Rent does not enter into cost'—is all too obvious. Few, however, would propose to tax consumer's surplus. Far

[1] *Co-ordination of the Laws of Distribution*, Sec. 7 (f).

from that, those who have built up a theory of taxation on any such notion have always been anxious to see rather that the damage done to consumer's surplus by tax measures is minimised.[1]

Directly analogous to, and in fact an extension of the consumer's surplus concept, is the idea of labourer's surplus and saver's surplus *in terms of satisfaction* secured in the act of exchange between real income and their respective sacrifices.[2] As consumer one gets a surplus because the 'total utility' of the commodity that he secures is greater than its total exchange value. As a worker one gets a surplus because the 'total disutility' of the work done is less than the total exchange value of his work. Similarly the saver gets a surplus because the 'total disutility' of his 'waiting' or 'abstinence' is less than the total exchange value of his capital. All these surpluses are strictly subjective because, as represented, they depend entirely upon indivi-

[1] The terminology originally used by Marshall to describe the gains from exchange had perhaps something to do with the prevalent notion that the two concepts referred to were analogous and were aspects of the same problem; although Marshall himself was well aware of the superficial character of the analogy. In his *Pure Theory of Domestic Values* and also in the first three editions of his *Principles* Marshall designates the surplus from exchange as 'Consumers' Rent'. Yet he emphasizes that "the analogy between the two theories of landlords' rent and consumers' rent, though close so far as it goes, does not extend far." (*Pure Theory*, London School Reprint, p. 23).

Jannaconne suggests (op. cit., p. 64 footnote) that the substitution of "consumers' surplus" for "consumers' rent" was meant to denote an analogy to the expression "Producers' Surplus" which designates the general character of the net benefit of production of which rent is a particular case. But this is wrong, as Marshall's own assertion would show. Speaking on Producers' Surplus he says: "This Surplus is a different thing from a consumer's surplus, which is the excess of the money value to a person of the satisfactions which he derives from a thing over the price which he pays for it". (*Principles*, 2nd ed., p. 454).

[2] Cf. Marshall, *Principles*, App. K; also Clark, *The Distribution of Wealth*, Chap. XXIV.

dual character and sensibility. They all represent surplus gains from Exchange.

The identity of character in them can be better exhibited by reducing the supply curve of work or abstinence against income in terms of the demand curve of income against work or abstinence. By so doing the so-called worker's surplus and saver's surplus might be shown to be nothing else than a sort of consumer's surplus. This reversal in the point of view has important analytical convenience. In the case of labour, for example, as Prof. Robbins has shown, the effect produced by a variation in the rate of wages on the willingness of workers to work can be most clearly shown with reference to the elastictiy of demand for income in terms of effort,—which elasticity may be greater or less than unity ; so that the belief held in some quarters that a tax on wages would be necessarily followed by an inclination on the part of workers to put forth more labour has no logical foundation.[1]

In the present context it has been necessary to indicate that the observations that will be made on consumer's surplus will apply equally to the so-called worker's surplus and saver's surplus, and further to emphasize the essential harmony between these surpluses. These 'subjective' surpluses, as they may be called,—all of them are ultimately based on the assumption of diminishing marginal utility of income. If the worker or the saver contributes to the production of real income, he is also a consumer of real income. In the case of the worker, his labour ultimately reduces itself to a question of an exchange of leisure for real income. In the case of the saver, his 'waiting' is a question of the exchange of present for future real income. In fact, in the

[1]Lionel Robbins, On the Elasticity of Demand for Income in Terms of Effort, *Economica*, June 1930. See also Dalton, *Public Finance* (sixth edition), Chap. X.

ultimate analysis, even an ordinary money-commodity transaction resolves itself into an exchange of subjective sacrifices. Although certain transactions manifest themselves as an exchange between money and concrete goods while certain others manifest themselves as an exchange between money and labour or other sacrifices, yet all these can be traced ultimately to an exchange between 'utilities' —differently estimated by different economic subjects. A buyer or a consumer in the ordinary sense holds money and buys goods ; but behind the money which he holds is some kind of labour or other sacrifices. The consumer's own resources are an embodiment of labour or other sacrifices and are themselves the result of an exchange. There is thus no inherent disharmony between consumer's surplus and the so-called worker's surplus or saver's surplus. This can be more clearly illustrated by an hypothetical case of a commodity produced by labour alone. From the point of view of the labourer (producer in the ordinary sense), production is but an exchange of the commodity in question against leisure ; and higher or lower cost merely indicates more or less leisure sacrificed for its production. In this sense the labourer himself is a consumer of the commodity in question in so far as he holds 'time' as an economic good. Any variation in the cost through taxes or bounties, for example, therefore reacts on the labourer producing the commodity and its so-called consumer in the same direction. If this case is typical of the complex system of production and exchange as we know it, then an important truth comes out, namely that there is normally no clash of interest between producers and consumers.[1]

[1] This simple consideration should be sufficient to indicate the essential weakness of the 'surplus theory of prosperity' (as one might call it) built up by the late Prof. Patten. Apart from the semi-metaphysical incursions into 'subjective cost as falling under dynamic economics, into

3. The idea of consumer's surplus is as old as economic speculation. It is implicit in all those early speculations that relate to a distinction between 'value in use' and 'value in exchange'; for, what after all is consumer's surplus but an estimate of the amount by which 'value in use' exceeds 'value in exchange'? "The things which have the greatest value in use have frequently little or no value in exchange; and, on the contrary, those which have the greatest value in exchange have frequently little or no value in use." This remark, for example, with which Adam Smith introduces his theory of value[1] is only couched in a technical language when it is said that "we cannot trust the marginal utility of a commodity to indicate its total utility."[2] Adam Smith, however, did not pursue the matter. As is well-known, without stopping to solve this riddle of value he went straight to investigating the problem of objective exchange value; the subjective aspect dropped out of the picture.

Ricardo had some shrewd observations on this question. As Marshall points out, in his discussion of the difference

'pleasure economy' and 'pain economy', into a 'total value' of goods as in general increasing "with the growth in the variety of consumption" and so on and so forth,—ideas to which no precise meaning can be attached,—it is not true, as Patten would have us believe, that there should exist an inverse relation between consumer's surplus and labourer's surplus, nor either that consumer's surplus, to use the term in any intelligible sense, should fall off with an increased diversification of consumer's goods. Patten's ideas which are criticised here are to be found in his *The Theory of Prosperity* and also in his Essay on the Theory of Economic Dynamics in Tugwell's Edition of his *Essays in Economic Theory*.

[1] The latter portion of the statement should not be taken too literally. Manifestly a thing cannot have a value in exchange if it has 'no value in use'. Perhaps Adam Smith was using the term 'value in use' in a normative sense, having in mind the idea of an ethical standard,—a vague and uncertain concept which has no place in Economics. (Cf. J. S. Mill, *Principles*, Bk. III, Chap. I, Sec. 2). But more probably the significant point for Adam Smith was the emphasis of a contrast.

[2] Marshall, *Principles* (8th ed.), p. 129.

between 'Value,' and 'Riches,' Ricardo was "feeling his way towards the distinction between marginal and total utility."[1] In fact, not only was Ricardo aware of the importance of the distinction between total utility and marginal utility and of the fact that the former decreases while the latter increases on account of a check to the supply of a commodity, he was further conscious of the 'subjective' character of utility and of the difficulty of its measurement. "Value in use", he says, "cannot be measured by any known standard; it is differently estimated by different persons."[2]

Yet these were hazy ideas not meant to form a technique of economic analysis. The earliest attempt to use the notion of surplus utility as an instrument of economic analysis was made by that engineer-economist of France, Dupuit,[3] whose name goes with Lloyd and Gossen as one of the early discoverers of the principle of marginal utility prior to Jevons and Menger, and with Cournot as one of the earliest to apply the mathematical method to economic analysis. Dupuit's theory, although crude, is in many respects an anticipation of Marshall's theory of consumer's surplus, and as such deserves a somewhat detailed description.

As one connected with public works, Dupuit got interested in the problem of price policy in regard to public utility services. The price of such services, he thought, should be fixed so as to maximise the benefit of the society rather

[1] *Op. cit.*, p. 814.

[2] *Principles of Political Economy and Taxation* (ed. Gonner), p. 420.

[3] Dupuit's essays *De la mesure de l' utilité des travaux publics* (1844) and *De l'influence des péages sur l'utilité des voies de communication* (1849) in which his main contributions on utility are embodied, were originally published in the *Annals des ponts et chausées*. These essays together with some other contributions have recently been reprinted in a volume entitled *De l'utilité et de sa mesure* (Torino, La Reforma Sociale) with a preface by Luigi Einaudi and an Introduction by Mario de Barnardi.

than with reference to maximum money return. This led him into an attempt to assess the benefit or utility which a society derived from the consumption of goods and services. He discovered that the utility of a thing differed with regard to different consumers and even to the same consumer with regard to different quantities consumed. As one consumes more and more of a commodity the utility per unit diminishes,[1] so that the market price (prix vénal) of a commodity cannot be made an index of its 'absolute utility'.[2] He thus found fault with J. B. Say's measure of utility, as the latter, not comprehending the phenomenon of a gradation of utility, had taken price to be "the measure of the value of things and their value the measure of their utility". On the basis of the principle of diminishing utility Dupuit showed how the price that a consumer actually pays for an object differs from the price which he would be willling to pay. The price actually paid depends on marginal utility,[3] and as the consumer would be willing to pay higher and higher prices had the quantity been smaller and smaller, the aggregate demand price exceeds the aggregate actual price. Herein is the source of what he calls 'relative utility', a surplus utility which a consumer secures from exchange. The 'absolute utility' that one derives from the consumption of a good is measured by the maximum price which he is disposed to

[1] In the text the proposition is put in an inverted form. Like Cournot, Dupuit treats the quantity consumed as a function of price; and he shows how the utility of an object diminishes less and less rapidly as the price rises and increases more and more rapidly as the price falls. "....à mesure de le prix d'un objet augmente, l'utilité diminue, mais de moins en moins repidement, et que quand ce prix diminue, elle augmente, au contraire, de plus en plus rapidement...." (*op. cit.*, p. 63). The procedure differs only in form from that which has been made familiar by Jevons who treats utility as a function of quantity.

[2] Dupuit's 'absolute utility' corresponds to Jevon's 'total utility'.

[3] Dupuit did not use the term 'marginal'; but it is implicit in his argument.

offer in order to obtain it. The excess of this price over that which he is obliged to pay is 'relative utility'.[1] Dupuit then proceeds to apply this notion for examining how the utility derived from the consumption of a good is affected by measures calculated to raise or lower its price. A tax on a commodity causing a rise in its price brings about a net loss of utility, and a measure calculated to lower price, on the other hand, brings about a net gain. If, for example, the cost of production of a particular commodity is 20fr. per unit and the utilities of different units of the commodity, as represented in terms of francs are arranged in order thus,

30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20fr.,

the corresponding relative utilities (absolute utility *minus* price) would be

10, 9, 8, 7, 6, 5, 4, 3, 2, 1, 0.

Now, if a tax is imposed to the extent of 5fr. per unit, then the new relative utilities (as diminished by 5 fr.) would be

5, 4, 3, 2, 1, 0.

So far the loss is counterbalanced by tax receipts. The net loss is due to the fact that some portion is not consumed at all on account of the rise in price, and the relative utilities of that portion in this particular case are

4, 3, 2, 1, 0.

By a parity of reasoning,—as regards the inverse hypothesis, it is shown that a diminution of price causes a net gain of utility owing to an increase of consumption whose relative utilities may be represented by

5, 4, 3, 2, 1, 0.[2]

[1] '....en général l'utilité relative ou définitive d'un produit a pour expression la différence entre le sacrifice que l'acquéreur consentirait à faire pour se le procurer et le prix d'acquisition qu'il est obligé de donner en échange." (*op. cit.*, pp. 40-41).

[2] This proposition, however, does not hold unambiguously for, if the diminution of price is due to a bounty per unit of the commodity, then the gain of the consumers is more than offset by the expenditure of the State.

But how are the utilities to be actually measured ? How are we to know what price consumers would be willing to pay for a commodity ? Dupuit attempts a precise calculation of this absolute utility of an object pertaining not only to an individual but also to a society. In order to follow his reasoning it would be convenient to refer to his geometrical construction which, in fact, apart from the crudity which will be noticed in a moment, forms one of the earliest mathematical demonstrations of the principle of diminishing utility. In the following figure,[1] ON repre-

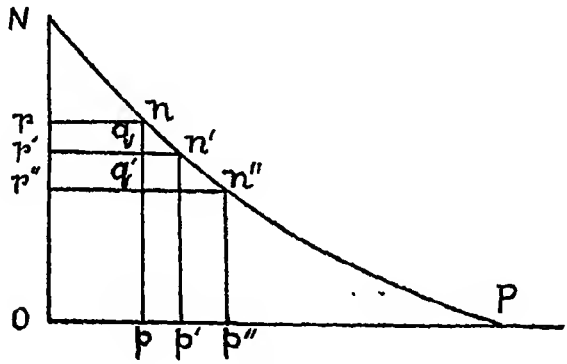

Fig. 3.

sents the quantity of a commodity consumed in the market when the price is nil, and OP the price when consumption is nil. In between them the ordinates denote the quantities consumed at different prices which are represented on the abscissa. If the cost of production (and hence the price) of the commodity is Op, the quantity consumed is np ; at the price Op', the quantity consumed is n'p' ; and so on. The curve, Nnn'n''P, thus formed is, as Dupuit calls it, the Curve of Consumption. For quantity, np, the utility per unit (as

[1] It is a combination of Fig. 1 and Fig. 3 of Dupuit with certain unnecessary details eliminated.

measurecd by price) is at least Op, and for n'p' the utility per unit is at least Op'. In np, therefore, there is some quantity, np—n'p'=nq=rr', the utility of which can be represented by rnn'r', for it contains one unit having at least Op' as its utility, the others having a utility less than Op'. Similarly, for r'r'', the utility is r'n'n''r''; and so on,—so that for quantity, np, of the commodity the absolute utility is rnPO. If the price per unit is Op, then the total sacrifice made by the consumers for securing the said amount is Ornp. The relative utility or the surplus is then npP. In actual practice, the points, n, n', n'' etc. on the curve of consumption, that is to say, the utilities of different amounts of a commodity, can be discovered, according to Dupuit, by taking an initial point and making continuous (small) alterations in the price throughout and then observing the effects on the amount consumed. The curve of consumption is then employed to examine the effect produced by a tax or a bounty on consumers' benefits. If, for example, a tax is imposed on the commodity so as to raise its price from Op to Op', the net loss of utility is nqn'; with a tax amounting to pp'', the net loss of utility is nq'n''; and so on.[1]

From all this Dupuit concludes that a State monopolist seeking to cover the necessary expenses and having an eye to the maximum benefit of the society would charge a different scale of price for the use of canals and railways from a monopolistic company seeking to maximise its own revenue.

There is also a clear reference to consumer's surplus in Jevons' discussion of *The Gain by Exchange*, where he criticises Mill's assumption of the terms of exchange as the sole determinant of gain from international trade.[2] "In estimating the benefit which a consumer derives from a

[1] It should be noted that Dupuit assumes that price rises by no more and no less than the amount of the tax.

[2] Cf. *The Theory of Political Economy*, pp. 142-46.

commodity, it is the total utility which must be taken as the measure, not the final degree of utility on which the terms of exchange depends".[1] The gain from exchange or surplus utility is easily deduced from his 'space-representation' of utility which is essentially similar to Dupuit's representation. Moreover, the utility curve employed by Jevons is not confined to the behaviour of an individual; it is also extended to the bahaviour of the market. Like Dupuit, Jevons also uses the demand curve of a 'trading body' as synonymous with their utility curve.

But, although Jevons imagined the gain from exchange in terms of consumer's surplus, he, unlike Dupuit, did not attempt to measure it. Jevons realised in particular that "the fact that we can most imperfectly estimate the total utility of any commodity should prevent us from attempting to measure the benefit of any trade."[2] To be able to measure it would be 'interesting'; but he found that it would be an impossible adventure. The notion of consumer's surplus does not therefore play any very important role in Jevons' scheme. The rectangles that figure in his 'space representations' have no more function to do beyond suggesting a 'variation of utility' with each increment of the supply of a commodity.

[1] *Ibid.*, p. 143. Elsewhere he expresses it much more explicitly. "The price at which a thing is bought is no necessary indication of its utility. Half an ounce of quinine may save a person's life, and, however rich he may be, he will not, in ordinary circumstances, pay more than a few shillings for it. But the benefit is immeasurably great. A book collector meeting with a rare and long sought edition may pay a price decidedly beneficial to the bookseller and yet may esteem himself fortunate. Its utility or interest for him may be so great that he would have paid far more if requisite. Now, what happens palpably in these cases happens also in every ordinary case of exchange." *Principles of Economics*, ed. Henry Higgs, pp 4-5.

[2] *Theory*, p. 145.

4. Now, although these demonstrations point to an important aspect of the gain from exchange, yet the manner in which it is worked out is open to serious objections.

One obvious comment on the way Dupuit and Jevons demonstrate the principle of diminishing utility is that it suggests an imputation of absolute degree of utility to different units of a commodity,—which is surely misleading. Even as a formal explanation of the variation of utility of an object Jevons' space-representation is defective. The utility of an object by itself is not significant in Economics. Like value and cost it should be taken in a relative sense. The behaviour of an individual at any particular moment suggests a scale of preference on which one thing stands higher than another; and that is all that utility analysis is entitled to assume. It is misleading to speak of "the utility attaching to any particular portion" of a commodity, as does Jevons, and as is implicit in Dupuit's argument. Separate units of a commodity do not appear on the market with specific labels of utility tied round their necks! And yet this is an easy interpretation that may be given to their way of putting things.

The limitations of Dupuit's money measure of relative utility will be more evident as we come to discuss in the next chapter Marshall's theory of consumer's surplus. But certain points may be noted here.

We have seen that Dupuit derives surplus utility from a downward-sloping demand curve,—the curve of consumption, as he calls it. This is on the assumption that the demand curve represents also the utility relation of goods exchanged. In fact, however, the demand curve indicates only a price-quantity relation and, in so far as it relates to the market which consists of heterogeneous consumers, it has no psychic significance. Consumer's surplus is essentially a subjective concept and relates to an individual.

And even from the point of view of an individual consumer, the demand curve is not a true index of utility. The utility function is more complex than appears in Dupuit's theorem. The maximum monetary sacrifice which a consumer would be willing to make for securing a commodity is a function of a number of variables. Primarily, no doubt, it is a function of the quantity of that commodity ;- but it is also a function of the prices and quantities of other commodities. In other words, the quantity of a commodity demanded by an individual buyer depends not only upon the price of the commodity but also upon the income of the buyer. Except on certain more hypothetical assumptions which will be noticed in the next chapter, it is therefore not possible to derive a measure of the surplus utility secured from exchange even from the point of view of an individual buyer.

As soon, however, as one leaves off the individual and comes to the condition of the market, the surplus referred to ceases to be 'subjective', and the demand curve becomes less useful as a guide to the study of surplus utility. The demand curve which reflects the price-quantity relation in the market is deduced from—indeed it is the resultant of—subjective individual preferences based, let us say, on utility, —but these preferences and utilities of individuals lie buried in the objective relationship manifested in the market and cannot be derived therefrom.

As has been pointed out, both Dupuit and Jevons made this confusion between demand price and utility,—forgetting that the utility analysis, while it may be legitimate in the case of exchange concerning an individual, cannot be extended directly to a group unless the group consists of homogeneous units. In the case of Jevons this confusion seems all the more striking when one finds him emphasizing in other contexts the subjective character of utility. In fact Jevons made it clear just at the beginning of his discourse that it

is not possible to "compare the feeling in one mind with that in another",—that "the weighing of motives must always be confined to the bosom of the individual".[1] But his hedonistic bias led him astray. Much too anxious for exhibiting the process of a balancing of 'pleasure and pain', he could not persuade himself to relinquish his favourite maxim of utility, and instead of starting with an individual (for whom he found the utility function to be discontinuous) he went straight to the circumstances of the market and took the 'average' to stand for the 'aggregate' as the basis of his analysis. The result has been that his analysis shirked and did not solve the problem of market equilibrium.

The objections to Dupuit's measure of surplus utility noted here had already been made by Walras[2] and have since been reiterated in refutation of Marshall's doctrine of consumer's surplus. But Marshall formulated his doctrine with full knowledge of these difficulties, and from the very beginning hedged it in with important assumptions which, as the history of the controversy over consumer's surplus shows, have often been overlooked.

Marshall's theory of consumer's surplus will be examined in the next chapter in the course of which it will be found that as it runs in the abstract his formulation is free from the limitations that have been seen to attach to Dupuit's theory.

[1] *Theory* 2nd edition p 14

[2] *Éléments* (3rd edition), pp. 442-445.

CHAPTER VII

MARSHALL ON CONSUMER'S SURPLUS

1. The concept of consumer's surplus occupies a prominent place in Marshall's system. It would perhaps be wrong to suggest that it forms any of those essential posts on which his theoretical structure rests. Nevertheless, it is a construction which provides that peculiar tone which distinguishes the 'school' of Marshall from other schools of economic thought. It is a construction which provides a basis of Welfare Economics. Marshall was not content with merely giving a description of a tendency to equilibrium of economic forces, which is avowedly the task of positive economics, but he paused to consider how far the equilibrium in question was consistent with maximum social welfare. In fact all his studies relating to the mutual adjustment of demand and supply (wants and activities) seem to have been subsidiary to a more fundamental problem of social welfare. He analyses the mechanism of price through which the forces of demand and supply are brought into equilibrium; but these analyses prepare, as it were, the ground for an attack on an ultimate problem of maximum satisfaction. And in the chapter on *The Doctrine of Maximum Satisfaction*[1], to which his previous analyses lead up, the apparatus of consumer's surplus is invoked for a demonstration challenging the complacent attitude of 'harmony' economists who saw in the natural working of economic forces a 'divine order' leading on to the maximum good of mankind, and who thus contemplated the position of equilibrium brought about by

[1] *Principles* (8th ed.), Bk. V, Chap. XIII.

the free play of natural forces as the position of maximum satisfaction.

Marshall originally expounded the doctrine of consumer's surplus in his *Pure Theory of Domestic Values*,—a paper which was printed for private circulation as early as 1879.[1] The account of the doctrine with which the readers of his *Principles* are familiar is already there in this paper. The elaboration that is made in the later editions of the book relates mainly to making more explicit the assumptions on which it rests. The satisfaction—so runs the argument—which a consumer derives from the purchase of a commodity exceeds "the value to him of the money which he pays" for it. The money that one expends on a commodity secures for him a greater satisfaction than he could obtain by expending that money on any other things; for, if any other mode of expenditure had seemed preferable, he would have chosen it. The 'economic measure' of this excess of satisfaction is consumer's surplus. Marshall derives this economic measure of the surplus from the 'demand-schedule' (or demand curve) of an individual which he assumes to indicate the comparative utility of the good secured and money sacrificed,—money representing 'other goods' which are displaced in the

[1] This paper along with his *Pure Theory of Foreign Trade* now forms No 1 in the series of Reprints of 'Scarce Tracts' in Economic and Political Science, London School of Economics. It should be noted that although, as we have seen, the doctrine of consumer's surplus took shape as early as 1844 in the hands of Dupuit, and is also a natural deduction from the 'space-representations' of Jevons, neither of them did actually influence Marshall in any way. Marshall deduced the idea from 'a study of the mathematical aspects of demand and utility under the influence of Cournot, von Thunen and Bentham'. (See *Principles*, 2nd edition, p. 184n). It seems Marshall had not read Dupuit when he first developed the idea; for neither in the *Pure Theory* nor in the first edition of the *Principles* is there any reference to Dupuit, although in later editions of the *Principles* Dupuit's name is mentioned more than once. Dupuit's work in fact had sunk into oblivion from which it was rescued by Jevons. See Preface to the second edition of Jevons' *Theory of Political Economy* (1879).

act of exchange. The surplus secured by a consumer from the purchase of a good is shown to be equal to the excess of the price he would be willing to pay—as reflected in his demand curve—over the price which he actually does pay.

Primarily as an abstract concept it is confined to the behaviour of an individual. But as an instrument for guiding practical policy Marshall suggests an extension of the principle to the behaviour of the market, too. And there he refers to the statistics of trade and deduces a surplus on the basis of the market demand curve.

We shall here concern ourselves mainly with considering the legitimacy of the concept as such relating to an individual in his exchange of goods and services. But we shall also have occasion to notice, *en passant*, the difficulties connected with the practical application of the concept in the contexts suggested by Marshall.

2. At first glance it seems as though the doctrine of consumer's surplus enunciated by Marshall is another name for Dupuit's Relative Utility. For, indeed, both are counting on an excess of what one would be willing to pay for a commodity over what he actually does pay. Moreover, Marshall, too, although he starts with an individual's preference, ultimately relies, as Dupuit does, on a downward sloping market demand curve for the deduction of consumers' surplus. Indeed even the use which Marshall makes of the concept is largely Dupuitian.

Yet, details of formulation apart, even in point of definition there is a subtle difference between the two which deserves attention. Whatever might have been in Dupuit's mind, the way he defines his surplus gives one an impression that after all what he is seeking for is an objective entity. According to Dupuit, the excess of the maximum price which a consumer is disposed to pay for a good over the actual market price *is* the surplus. The difference, if any, between the two

prices is simply a sum of money, and this sum is taken to be a measure of the extra satisfaction which a consumer secures from purchasing the good. In Marshall's definition, on the other hand, the subjective element is brought out more vividly. The surplus is the excess of satisfaction which the consumer secures from his economic environment, this excess being the difference between the utility of the good acquired and the utility of the good (or a collection of goods represented by money) sacrificed. This subjective element is primary in Marshall's representation of consumer's surplus, its expression in money terms is secondary.

This deserves particular attention in as much as a failure to observe the emphasis suggested here has led to an unnecessary misunderstanding and has been responsible for one entirely unwarranted piece of criticism of Marshall's doctrine. Taking consumer's surplus to mean simply a difference between the potential price and the actual price Ulisse Gobbi, for instance, has argued that in the ultimate analysis this surplus must necessarily be reduced to zero.[1] When account is taken of the totality of purchases that a consumer makes, the price which he would be willing to pay just coincides with the price which he actually pays; because both are limited by the amount of money that he has command over, that is to say, by his income. If an individual starts with a given income he may be supposed to be willing to spend the whole of it on one good. As, however, he secures the good for a smaller sum, he turns to a second good, this time offering only what he is left with after the purchase of the first good. If again there arises a margin he turns to a third good, and so on,—the margin between the offer price and the actual price becoming narrower as the series of actions is extended, until it vanishes as he completes his final purchase.

[1] Gobbi, Un Preteso Difetto delle Imposte sui Consumi, *Giornale degli Economisti*, 1904

This is true. If the surplus had been a matter of difference simply between two sums of money, as Dupuit's loose way of defining it seems to suggest, we should indeed have found it reduced to zero, having regard to the total position of the consumer in respect of his purchases.

Hobson makes a queer case for consumer's surplus.[1] While urging that it is merely an expenditure-income relation in price terms and that it must be reckoned with reference to the total purchase and the total income of a consumer, he yet finds a consumer's surplus in the volume of saving which the consumer is enabled to make under a given situation. No such surplus, he argues, can arise when the whole income is spent on current consumption. Only when a portion of one's income is saved can there arise a consumer's surplus. "If I do not spend, but save, a portion of my income, that saving rightly appears as consumer's rent, even when the totality of purchases is set against the total income, for I would have sacrificed the whole of saving rather than have dispensed with a necessary and some of it rather than go without a convenience".[2] But saving may—as it usually does—represent a desire for future income in which case obviously it is only another name for *spending* for future goods. Even if, as one might argue, it represents merely a preference for cash holding, one fails to see why this motive itself should not be included within the consumer's scheme of choice. Surely the truth remains that what a man earns he consumes.

Marshall's doctrine of consumer's surplus, however, has a different significance. Clearly what Marshall seeks to emphasize is that with given resources an individual may find himself better off under a more favourable economic environment. The fact that he is enabled, as in Gobbi's

[1] *Economics of Distribution*, Chap. 11.

[2] *Op. cit.*, p.47.

illustration, to extend his purchases with a given income to a variety of goods is itself a benefit which comes under consumer's surplus. If consumer's surplus evolves *within* a particular line of purchase, it is *outside* of that line that it realises itself.

If this interpretation is correct, then it becomes at once clear how inconceivable is the idea of summating the consumer's surplus derived from the whole complex of goods that the consumer purchases. An aggregate of consumer's surplus is surely a myth.

It should, however, be pointed out that Marshall, repudiating as he does any suggestion that his doctrine might appear to convey of a summation of consumer's surplus, would yet be inclined to believe that it is 'theoretically feasible'. He seems to find difficulty in the task only in so far as the commodities that come under exchange are related to one another,—in so far, that is to say, as the condition of supply of one affects the utility of another.[1] But once the relative character of utility and consumer's surplus is recognised it becomes clear that the problem is not only a practical impossibility, it is a theoretical absurdity. Even supposing that in the last analysis it is a question of a comparison of the pecuniary with the non-pecuniary use of one's resources—and this is perhaps what Marshall meant in this context, although such a standpoint at once precludes the possibility of any money measurement of the surplus,—there is no escaping the fact that the non-pecuniary use itself is an object of exchange. In fact, the idea of a 'sum' of consumer's surplus begs the assumption—long since discredited among the psychologists themselves—of pleasure and pain as two antithetical feelings in man which can be measured and compared.

[1] *Principles* (8th ed.), footnote on p. 131. The point was first noted in the third edition, p. 207n.

3. Gobbi further asserts that even in respect of a single-line purchase the consumer's surplus is zero. "The difference between that which is spent and the maximum which one would be willing to spend is zero not only for a complex of goods but also for a single object so long as a diminution of price corresponds to an increase of consumption".[1] So far as 'economic utility' of a good is concerned, it is argued, the marginal utility multiplied by the quantity *is* its total utility, for, the price per unit which a person is willing to pay for it already reflects the utilities of all the units contained in the amount purchased. It is the total demand at each particular price that is significant for economic calculation, and, so long as a good is perfectly divisible, at the point of equilibrium the potential price coincides with the actual price and there is no scope for the emergence of a surplus.

A similar line of argument is advanced also by Cannan and others in refutation of the theory of consumer's surplus.[2] "Ten times the price of one-tenth", thus runs the argument of Cannan, "is the proper price of the whole with all its utility attached, and before people were confused by misleading 'space-representations' nobody ever imagined that he was getting something for nothing when he paid the whole market price."[3]

As has been noted, this is an obvious criticism against Jevons' space-representations. Those rectangular representations seem to imply as though the different units that are

[1] Sulla Misura del Vantaggio che il Consumatore ritrae da uno Scambio, *Giornale degli Economisti*, Feb. 1916.

[2] See Cannan, "Total Utility" and "Consumer's Surplus", *Economica*, Feb. 1924. See also Knight, *Risk, Uncertainty and Profit*, footnote on pp. 71-72; Economic Psychology and Value, *Quarterly Journal of Economics*, 1925, p. 374n; Harry E. Miller, Utility Curves, Total Utility and Consumer's Surplus, *Quarterly Journal of Economics*, Feb. 1927.

[3] *Op. cit.*, p. 22.

placed on the market have different utilities attached to them. In fact, however, when they are placed on the market there is no distinguishing one from the other unit, for, each of the units being perfect substitutes must afford the consumer the same amount of satisfaction. But Marshall's 'utility schedule' is free from this ambiguity. In the example of tea that is given to illustrate the principle, what is referred to is clearly a comparison between the utility of tea and the utility of money (money being taken to represent other goods) under different hypotheses concerning the supply of tea,—a comparison showing how at each successive point the utility of tea stands in relation to that of other goods, and what the position of the economic subject would be in view of each successive decrement (or increment) in the supply of tea. And, of course, taking, for example, a limiting hypothesis that tea is non-existent and the consumer is having to spend the money hitherto spent on tea on other goods it would be seen that he gets worse off. All that Marshall's 'utility schedule' should be taken to represent is the loss of utility contingent on the withdrawal of successive units of a good (or gain of utility on additions thereof) as expressed by the consumer's willingness to part with other goods. And, in fact, the chief concern of Marshall in introducing the apparatus of consumer's surplus was to compare the loss or gain of utility on account of price changes as between different lines of purchase.

But is any precise comparison of that sort possible?

4. This brings us to the assumptions which Marshall makes in order to effect such comparison.

In the first place, Marshall assumes that the marginal utility of money to a consumer is constant throughout the process of exchange. This assumption indeed runs through the whole of Marshall's analysis of value; for, it is with the aid of this assumption that he could solve the problem of

determinateness of market equilibrium.[1] On the doctrine of consumer's surplus the bearing of this assumption is that only by taking the marginal utility of one of the goods exchanged to be constant is it possible to talk in precise terms of any comparison between the utility secured in an exchange and the utility sacrificed.[2]

In the second place, Marshall considers the utility of a good as dependent on the supply of that good alone. Complementarity of goods, so far as the valuation of goods of the first order is concerned, is to Marshall as it was to Mill a special case. To use a technical term, the good whose utility schedule is under consideration is an 'independent' good. This second assumption is already included within the first; for, if of two commodities entering into exchange one has a constant marginal utility then it directly follows that the commodities are independent. Yet in view of its importance in Marshall's analysis the assumption deserves explicit mention.[3]

[1] See Marshall, *Principles* (8th. ed.), Bk. V, Chap. II; also Appendix F.

[2] This assumption was not stated in *Pure Theory*. In the first edition of the *Principles* there is a passage which suggests it. "For one and the same time", it is stated, "his (the consumer's) material resources being unchanged, the marginal utility of money to him is a fixed quantity, so that the prices he is just willing to pay for two commodities are to one another in the same ratio as the utility of these two commodities." (Bk. III, Chap. II, pp. 155-156). But here what is guarded against is more particularly the variability of the marginal utility of money on account of changes in the wealth of the consumer over time. In the second edition (1891), however, the assumption is more explicitly stated. See footnote on p. 182. The whole thing was further cleared up in the controversy over the doctrine between Nicholson on the one hand and Edgeworth and Barone on the other. See Nicholson, *Principles of Political Economy* (1893), Vol. I, Chap. III, particularly *Note on Marshall's Treatment of Consumer's Rent* at the end of the chapter; Edgeworth, Professor Nicholson on Consumers' Rent, *Economic Journal*, 1894; Nicholson, The Measurement of Utility by Money, *Economic Journal*, 1894; Barone, Sulla "Consumers Rent" *Giornale degli Economisti*, 1894.

[3] In algebraic terms this assumption means that the utility function takes the form $U = f(x) + \phi(y) + \psi(z) + \ldots\ldots$ instead of the form $U = f(x, y, z \ldots)$, where U stands for utility and x, y, z etc. for the commodi-

Granted these assumptions, Marshall's construction stands. So far as any individual is concerned, if the marginal utility of the good that he is parting with is not affected by more or less of it that he possesses, and if the good that he is purchasing has a utility which is independent of other goods, then a comparison of the utility secured in the exchange and the utility sacrificed has a definite significance.

This can be demonstrated in more precise terms on Indifference Curves:

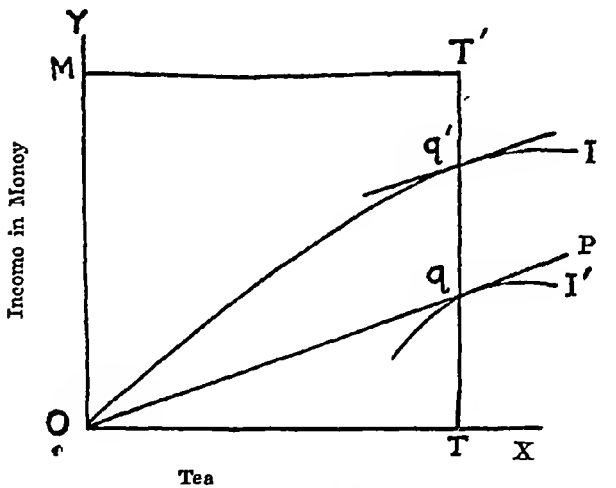

FIG. 4.

ties consumed by the individual. The latter form was first used by Edgeworth to indicate the general dependence of utility (Cf. *Mathematical Psychics*, p 20). Marshall, on the other hand, uses the other form. "Prof. Edgeworth's plan of representing U and V as general functions of x and y has great attractions to the mathematician; but it seems less adapted to express the everyday facts of economic life than that of regarding, as Jevons did, the marginal utilities of apples as functions of x simply." (*Principles*, Mathematical Appendix, Note XII). Now, once we assume constant marginal utility of money or 'other goods', the condition of independence of the good exchanged follows automatically. If M represents money and T represents tea, it is easy to see that the

Let us suppose (Fig. 4) that our consumer starts with OM quantity of money and comes to the market for purchasing tea. Money is measured along the vertical axis and tea is measured along the horizontal axis. If the marginal utility of money is constant, the rate of substitution between money and tea at vertically consecutive points on the indifference curves is the same, for that rate depends only upon the quantity of tea exchanged. Let OI be the zero indifference curve[1] which starts from the origin, and OP the price line touching a higher indifference curve I' at q; then TT', a line drawn perpendicular to OX and passing through q is the Expenditure Curve,[2] showing that any saving of money made in the transaction does not at all affect the consumer's demand for tea. If TT' cuts OI at q', then it is easy to see that qq' is a surplus which the consumer secures from the purchase of tea. The zero indifference curve, OI, indicates that the consumer would be willing to pay Tq' of money in exchange for OT of tea. At that rate the exchange would make him neither better off nor worse off.[3] But the price ratio in the market being what it is, he has to pay Tq of money for OT of tea, and none of the saving represented by qq', is spent on tea,[4] so that qq' signifies a

function expressing the utility of M and T must take the form $U = f(M) + \phi(T)$; for, if the marginal utility of M is constant it must be independent of the supply of T. As between themselves the commodities that are represented by M may be 'complementary', but with respect to T they are on the whole 'independent'.

[1] Cf. Bowley, *Mathematical Groundwork*, Chap. I.

[2] Cf. Hicks and Allen, A Reconsideration of the Theory of Value, *Economica*, Feb. 1934.

[3] OM (money) with which the consumer starts = T' q' (money) + MT' (tea) = T'q' (money) + OT (tea), or OT (tea) = OM (money) − T'q' (money) = TT' (money) − T'q' (money) = Tq' (money).

[4] We are assuming here that the consumer's indifference map is given and is not itself influenced by the market price. There may, however, be cases where the intensity of desire for a commodity is itself a function

definite surplus which can be compared with such surpluses elsewhere.[1]

5. All these are, however, arbitrary assumptions, and by making them, Marshall has left out of account elements which are not always of the second order of smalls.

It is not legitimate to take the marginal utility of money as constant. If we use money as *numeraire* in the Walrasian sense and take it as one of the commodities entering into the circle of exchange, then the arbitrariness of the assumption becomes at once obvious. If, on the other hand, money is taken as a representative of other goods, —an assumption which we have so far made—then the utility that we impute to it is derived from other goods and at each successive price-ratio on which the demand curve is based we have to take account of the satisfaction secured through an optimum distribition of money on other goods, and there also the marginal utility of other goods is certainly subject to variation. It should increase as more and more money is withdrawn from them. To make the point clear we may revert to the limiting case, mentioned in the last section, of a complete withdrawal of tea from the market. The money hitherto spent on tea will be spent on other goods, and if there is an optimum distribution of money among these goods, the difference between the utility destroyed on account of the withdrawal of tea and the utility obtained (the maximum that could be drawn from distributing the money released from tea among a variety of goods) will be much less than is shown in Marshall's construction.

of price. High price, for instance, may be a virtue raising a commodity in the estimation of its consumer. But these cases are rare and may legitimately be ruled out.

[1] For a similar demonstration, see Edgeworth, Review of Marshall's *Money, Credit and Commerce*, Economic Journal, 1923; also Hicks, *Value and Capital*, Note to Chap. II.

Now, of course, this limiting case is far from being typical of that which Marshall had in view. Marshall considers the case of a commodity on which the consumer spends a very small part of his income and enquires into the effect of a small change in its price. If indeed we are dealing with infinitesimals, a withdrawal of money from—or a transfer of money to—a large number of goods would affect but little the marginal utility of any one of them taken in isolation (if the good in question is 'independent'). But the limitation involved in the assumption is that we are leaving out of account a margin of error which, taken jointly, may assume significant proportions. And the limitation becomes more important when, as in real life, we take finite changes.

Now, once we allow for a variation in the marginal utility of money the demand curve becomes less straightforward than is shown in Marshall's construction. The different points on the consumer's demand schedule for a good cease to be co-existent; they relate to shifting hypotheses, so that a straight reading off of consumer's surplus from demand curve becomes impossible.

Similarly, taking two or more commodities it is found that their demand conditions in terms of money are not co-existent. The amount of money that one would be willing to pay for coal would depend not only upon his desire for coal but also upon what amount he has already spent on, say, bread. The consumer's surplus from coal, therefore, would be different on the assumption that bread has already been purchased from what it would be on the assumption that the consumer starts initially with coal. And so with bread. The surplus from one cannot in these circumstances be compared with the surplus from the other. The attempt to do so would look like starting to measure length by an elastic yardstick.

Further complications arise from the inter-dependence of the utility of goods. In the last analysis every good is related to every other good. The utility of one good is a function not of the quantity of that good alone, it is a function of the quantities of all goods in the market.[1] But even apart from this general dependence which Marshall rules out as being "less adapted to express the every-day facts of economic life", there is hardly any commodity which is not immediately related to some other commodity one way or the other. The very example of tea (or coal, as in earlier editions) that Marshall uses in enunciating his principle is typical of such dependence.

Marshall proposes to group such commodities under one common demand schedule.[2] But is that possible ? This raises the problem of defining a commodity which can only be touched upon here very briefly. In defining a commodity, are we to start from a particular kind of want and observe which goods actually cater to that want, or should we start from a concrete good consisting of perfectly substitutable units and estimate the importance of the uses to which it caters. The latter method surely appeals more to common-sense. But quite apart from that, the former method raises more fundamental difficulties. In the first place, it is difficult to define a want specifically. Without going so far as to suggest that human want is a unified whole,—that in the ultimate analysis it reduces itself into one supreme desire 'to live',—a philosophical complication which for our more modest purposes we can waive,—one feels doubtful about the efficacy of drawing a rigid boundary line between different *kinds* of wants. Secondly, the same concrete good may be used for different purposes. Thirdly,—

[1] Cf. P. N. Rosenstein-Rodan, La Complementarieta, *La Reforma Sociale*, 1933.

[2] *Op. cit.*, pp. 100n, 131n.

what is more important,—there is the difficulty of choosing units in terms of which the diverse elements should be compared so as to be fitted into a demand curve. If they are chosen in terms of relative valuation and substitutability at the margin, there is no knowing where to end. One would have to carry it to its logical conclusion and face the proposition that "if I am indifferent as to whether I have a wireless set for £ 10 or whether I have the satisfaction of saving ten Chinese children from starvation, the wireless set in London is the same quantity of the same commodity as £10 worth of rice in China".[1]

These limitations in the assumptions on which Marshall's construction is based suggest how, even if we proceed from the point of view of a single individual, it is not possible precisely to compare the consumer's surplus derived from one commodity with that derived from another. Yet they do not lead to a denial of the surplus as such. Consumer's surplus signifies a passage from a lower to a higher indifference curve which environment makes possible for an economic subject. The consumer in purchasing a commodity against one which he himself possesses aims at reaching the highest point on the 'hill' of preference which is consistent with his environment, and he achieves his aim by substituting that quantity of his own commodity for the one that he is purchasing at which their relative marginal utilities to him conform to the price ratio in the market. In other words, he stops at a point at which the price line is tangential to one of his indifference curves. This is precisely the condition in which his surplus is maximum. In fact this latter is one way of describing a partial equilibrium.[2]

[1] *Review of Economic Studies*, Vol. I, No. III, P. 167.

[2] See, *e.g.*, Fisher, *Elementary Principles of Economics*, p. 295; Bowley, Does Mathematical Analysis Explain? A Note on Consumer's Surplus, *Economica*, June 1924. In this article Prof. Bowley after stating how

6. It was, however, not as an aid to the explanation of equilibrium that the apparatus of consumer's surplus was introduced. Nor is it quite an effective tool for that purpose; for it breaks down as soon as more than two commodities are brought into the scheme with all their correlations taken account of.[1] On the other hand, Marshall applies his doctrine of consumer's surplus to considering the effects under different circumstances of taxes and bounties upon the 'economic welfare' of a community.

Leaving out of account the question of landlords' rent,[2] Marshall suggests that a tax on a commodity produced under diminishing returns may result in a net gain of utility; for, the gain of the State by way of tax receipts may more than offset the consequent loss of consumer's surplus; while by a parity of reasoning, a bounty on a commodity produced under increasing returns may result in a gain of consumer's surplus which would more than offset the State expenditure on the bounty. He argues moreover that, other things remaining the same, a tax on a commodity

in a single-line purchase the consumer maximises his surplus, proceeds to suggest that a general equilibirium with respect to the totality of purchases can also be explained in terms of surplus. Taking several commodities, X_1, X_2, X_3,..., with prices p_1, p_2, p_3..., if F (x_1 x_2 x_3..) is the greatest sum that the consumer would pay for a basket of x_1 units of X_1, x_2 units of X_2 etc., then, Prof. Bowley argues, the quantity to be maximised is $F(x_1\ x_2\ x_3....) - p_1\ x_1 - p_2\ x_2 - p_3\ x_3.....$ But it is difficult to see what the significance of this surplus is. If the two expressions refer to two sums of money, as evidently they do, then the 'resulting surplus', as we have seen, is zero.

[1] Cf. Fisher, *Mathematical Investigations in the Theory of Value and Riches*, Part II, Chap. IV.

[2] This reservation is important, and Marshalll disposes it of not without due regard to the limitation it involves. The case that he endeavours to make out in this connection is one of taxing an industry where diminishing returns act very sharply, and out of the proceeds of the tax subsidising an industry where, on the other hand, increasing returns operate sharply. If, then, the tax involves a fall in demand for land and

having an elastic demand destroys more of consumer's surplus than a tax which brings in the same revenue imposed upon a commodity having an inelastic demand.[1]

Further, in the theory of monopoly price the doctrine has been employed to demonstrate that a monopolist seeking to serve the interest of consumers would fix a lower price and offer for sale a larger amount than if he is solely interested in his own net revenue,—and that this tendency will be the greater the more eager he is to add to consumer's surplus. By an extension of this latter principle, it is even argued that there may be cases where an undertaking for which the supply curve is above the demand curve throughout its whole length,—which is therefore unremunerative in money terms at any price whatsoever,—may yet secure a net gain to the community, the consumer's surplus that results from it outweighing the loss in money.[2]

These practical counsels directed towards an increase of social welfare through State interference involve further complications arising from a lack of homogeneity in the groups of consumers affected by such interference. As we have seen while discussing Dupuit's 'curve of consumption', there lie concealed beneath the market demand curve a host of indifference maps relating to individuals having varying tastes and sensibilities. It is true that here also if we concentrate on one commodity and on the position of the consumers of that commodity we should find that they

a consequent loss of landlords' rent, that loss will be, to some extent at any rate, be counterbalanced by an increased damand for land coming from the subsidised industry. And the net loss of rent may be small as compared to the gain of consumers. See *Pure Theory*, Chap. 11, Sec. 5; *Principles* Bk. V, Chap. XIII, secs. 6-7, particularly the diagrammatic analysis in the footnote on p. 473.

[1] Cf. *Principles*, p. 467n.

[2] *Op. cit.*, Bk. V, Chap. XIV, Secs. 7-8.

attain a more and more preferred position as they rise higher up on the Integral Demand Curve, and *vice versa*. But the problem is one of estimating the extent of this preference and of comparing the position of one group with that of another when they are confronted with price changes in their respective lines of purchase.

Marshall was aware of this limitation arising out of the incomparability of inter-personal experiences of satisfaction. He noticed it even in his *Pure Theory*.[1] But he proposed "as a first approximation, to treat a pleasure that is worth a shilling to one man as equivalent to a pleasure that is worth a shilling to any other man". He suggests further that when we are dealing with groups of consumers among whom heterogeneous characters are distributed in normal proportions, personal peculiarities would cancel one another, and, on the average, the assumption would hold. But the statement of the limitation itself should be sufficient to show with what degree of caution the assumption has to be taken. With respect to ordinary necessaries of life it may be approximately true. But in a large number of cases it may prove to be the reverse of the truth. Any generalisation based on an assumption like this would surely be open to question. It is thus difficult to accept Marshall's position when he says, "The economic measures of various satisfactions can be represented in statistical tables ; and these may be used in establishing economic laws".[2]

Now, of course, the maximum of consumer's surplus that is seen to arise under free exchange in any given line of purchase is a relative maximum. The implication of a regime of free competition is that *given a particular economic*

[1] *Pure Theory of Domestic Values*, p. 22. Marshall even took Jevons to task because the latter had applied without qualification his maxim of utility to the case of a 'trading body'. See *Principles*, p.101n ; also p.818.

[2] *Pure Theory*, p. 22

environment the consumer adjusts his expenditure so as to reach what he considers to be his maximum position. There is nothing in it which ensures that the position reached is, absolutely speaking, a position of maximum welfare. Surely one cannot, in spite of Bastiat, comfort himself into a complacent belief in the beneficence of a 'natural order' like the poet who sang,

God's in his heaven—
All's right with the world !

Indeed there may be wide areas in which State interference would be necessary and beneficial. Some undertakings, if left to private enterprise, may be carried too far, absolutely speaking ; others may not be carried far enough. (The latter case is more likely under conditions of monopoly or imperfect competition). Yet in view of the limitations just considered any attempt to apply Marshall's construction of consumer's surplus with a view to discovering particular cases of disharmony is of doubtful efficacy.

In considering the effect on welfare of taxes and bounties, for example, one should take account of the character of consumers injured or benefited,—and, what is more important, of the repercussions of a shifting of resources which would be likely to follow these measures.

Take, for example, the principle of taxing commodities with a low elasticity of demand. In general, these are the commodities high or low price of which interests the poorer section of a community ; and the extent of loss that these people suffer on account of a high price of the necessaries of life cannot certainly be compared with the effect of a tax on a luxury good which is consumed by the richer section. This is an obvious difficulty. What is, however, very often overlooked in this connection is the effect of a shifting of consumers' resources. In the event of a tax upon—and a consequent rise in the price of —a commodity with a low

elasticity of demand, the total expenditure on it on the part of consumers becomes larger than before, and this extra amount is to be withdrawn from other lines. On the other hand, if the commodity taxed has a high elasticity of demand, the total expenditure on it becomes less than before, and the resources released are spent on other goods. When these repercussions are recognised one finds how illusive it is to read off from the ordinary unit-demand-curve the loss or gain of consumers' surplus on account of taxes and bounties.[1]

So far as the relevance of the concept to the theory of monopoly price is concerned, it is no doubt obvious that the monopolist in fixing the price of his commodity can—and sometimes does—take into account the benefit of consumers. But one can hardly get a precise idea about the 'total benefit curve' or the 'compromise benefit curve'[2] so as to be able to weigh the monopoly net revenue that is sacrificed against the consumers' gain that is secured. In view of this it is difficult to pronounce on the propriety of an undertaking for which the demand curve is throughout its whole course below the supply curve. It is true that if in such circumstances the State undertakes to bear the loss from out of general revenue, a commodity which is otherwise unremunerative is produced, and the consumers are benefited that way. But the benefit is secured at the expense of the

[1] It is sometimes argued that a tax on a commodity with inelastic demand is preferable to a tax on one having an elastic demand on the ground that the former affords the consumer greater freedom in the adjustment of his expenditure. But it is difficult to see why, if we start with a *given amount* of tax, its incidence, so far as the distribution of consumer's resources is concerned, should be at all different. Of course. the consumer, if he is rational, will distribute his new income (the original income *minus* the amount of the tax) in the same way in either case. Propositions of this sort seem to neglect the *general* effect of a tax on the distribution of resources among possible alternatives.

[2] Cf. Marshall, *Principles*, Bk. V, Chap. XIV.

tax-payers and at the expense of some resources which, but for this State intervention, would have been utilised in other ways. The same criticism applies to a similar theory first advanced by Dupuit and later on adopted by Edgeworth, namely, that it pays the society to allow a monopolist to appropriate consumers' surplus through price discrimination on the ground that such discrimination may enable the monopolist to produce a commodity which otherwise could not have been profitably produced. Here also it should be remembered that these resources are employable in other directions, and one has to weigh the benefits secured in the two cases before one can draw any significant conclusion.

CHAPTER VIII

"CONSUMER'S SURPLUS" IN INTERNATIONAL TRADE

1. As has been noted in Chapter VI, it was to estimate the gain resulting from international trade that one of the earliest applications of the principle of Consumer's Surplus was made. Since Jevons' attack on the classical doctrine of gains in international trade the role of consumers' benefits in the determination of such gains has assumed an importance.

One word of explanation why the problem of gains from international trade is separately treated is needed. It may be legitimately thought that international trade is but a species of exchange and that the problem of gain from international trade is but a special case of the general problem of gain resulting from division of labour and exchange. In the second place, it may be argued that the limitations already discussed of estimating consumer's surplus apply with added emphasis in the present context. For, a 'nation' is composed of heterogeneous individuals who are differently affected by alterations in the relative prices of goods ; and further, when it concerns large groups of individuals, the effect of alterations in the rate of interchange between any two commodities or groups of commodities upon the prices of other commodities becomes all the more important. Yet there is one respect in which the problem of international trade differs from the problem of distribution of income within a country. In the case of internal trade, if we are equally interested in the demand conditions of all the parties to exchange, our analysis relates simply to the condition of maximum *total* product. In the same way in the case

of international trade, if we regard the interest not only of one particular country but the interests of all parties concerned, the maximum *total* income to the entire region under trade relations is in question ; and the presumption in favour of free trade leading to the most intense international division of labour holds unambiguously.[1] If, on the other hand, regard is had to the interest of only one of the parties,—the 'home' country,—then we are concerned with the gain from trade of that country even though it may be at the expense of other countries. Economic analysis should enable us to see if, to use Mill's language, "any country, by its own legislative policy, can engross to itself a larger share of the benefits of foreign commerce, than would fall to it in the natural or spontaneous course of trade".[2] And in fact it appears that one of the major justifications of a separate study of the theory of international trade is to discover how far, if at all, there is a clash between nationalism and internationalism in Economics. For this, what we are required to know is not the absolute measure of the gain from trade but merely whether the gain is *greater* or *less* under different hypotheses concerning the rate of interchange between the export good and the import good.

2. In his classic chapter on International Values John Stuart Mill argues that the total gain from international trade depends upon the difference in the comparative costs of the goods exchanged in the countries in question, and that the proportion in which this total gain is distributed between the parties depends upon the rate of interchange, which again depends upon the play of reciprocal demand. The largest share of the gain is secured by those countries

[1] Let it be noted that it was in connection with the doctrine of free trade that the classical economists started the analysis of gains from international trade.

[2] *Unsettled Questions of Political Economy*, p. 21.

the demand outside for whose products is the greatest and whose demand for foreign products is the least. From this Mill concludes that the richest countries gain the least from foreign trade, since "having a greater demand for commodities generally they are likely to have a greater demand for foreign commodities and thus modify the terms of interchange to their own disadvantage".[1] Basing his conception of gain on the rate of interchange Mill further urges that it is possible for a country by means of taxes on imports or exports to turn the rate of interchange in such wise that it can "appropriate to itself, at the expense of foreigners, a larger share than would otherwise belong to it of the general productiveness of the labour and capital of the world".[2]

Jevons objects to this way of looking at the gain. Gain from exchange, he suggests, results from a difference between 'totalutility' received and 'total utility' parted with. Commerce is "the Exchange of the comparatively superfluous for the comparatively necessary", and any such exchange gives rise to a net gain of utility. In taking the rate of interchange as the basis for calculating this gain, Mill makes a confusion between 'total utility' and 'final utility', for, as he urges, the rate of interchange of which Mill speaks relates to the final utility, whereas what should be relevant in this context is 'total utility'.

Jevons thus criticises Mill's conclusion that a rich country derives less benefit from foreign trade as being the opposite of the truth. An intense demand for foreign product on the part of any country, far from being the cause of smaller gain, rather reflects a more favourable situation; for, although it might be accompanied with a seemingly unfavourable rate of interchange, yet the utility scale that lies behind it would register a benefit. A higher rate of inter-

[1] *Principles of Political Economy*, Bk. III, Chap. VIII, Sec. 8.
[2] *Op. cit.*, Bk. V, Chap. IV, Sec. 6.

change offered denotes either that the commodity imported is of a high degree utility or that the commodity exported against it is of a low degree utility. In either case the exchange yields a gain. Jevons illustrates his contention in the following manner.[1]

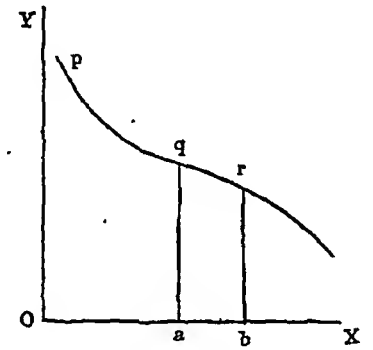

Fig. 5.

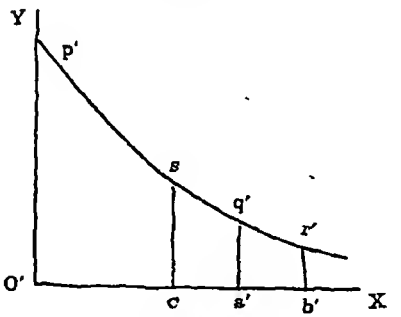

Fig. 5-A.

If the total utility of ob quantity of wool to, say, Australia, is represented by obrp (Fig. 5), and of o'b' of cotton goods by o'b'r'p' (Fig. 5-A), then an exchange of wool against cotton goods done by Australia at the rate, say, ab=o'a' gives a surplus of utility equal to o'a'q'p'—abrq, and as o'a'q'p' is a larger area than abrq (for, otherwise the exchange

[1] *The Theory of Political Economy* (4th ed.), p. 144.

would not be effected) it is a positive surplus. Even if the rate had been ab=o′c, apparently a less favourable rate for Australia, it would yet result in a net gain of utility, for the utility of o′c of cotton goods would have exceeded the utility of the wool parted with.

This analysis of Jevons is evidently based on what we have described as the principle of consumer's surplus. But Jevons considers the gain thus inferred as not susceptible of measurement and does not go beyond suggesting simply that as "no one will buy a thing unless he expects advantage from the purchase.....perfect freedom of exchange..... tends to the maximising of utility". He thus questions the validity of Mill's propositions regarding taxes on imports and exports.[1]

Now, clearly the above demonstration of Jevons is vague and cumbrous. In the first place, as we have abundantly shown earlier, no precise significance can be attached to the 'total utility' areas referred to in the diagram. Secondly, it does not give us any idea of the equilibrium position of exchange ; it does not tell us why ab is exchanged for o′a′ or o′c. Thirdly, in the context in which it has been used the demonstration seems to have missed its aim. It only establishes this, that even at a higher rate of interchange Australia secures *some* gain, and Mill would not deny that possibility. But even on Jevons' showing the magnitude of the gain appears to depend upon the rate of interchange. The area o′csp′ being less than the area o′a′q′p′, the gain that Australia secures from exchanging her wool for cotton goods diminishes as the rate becomes unfavourable to her.

In criticising Mill, what Jevons had to show (and what perhaps was in his mind) is that a lower rate of interchange

[1] *Op. cit.*, p. 146. "Customs duties may be requisite as a means of raising revenue, but the time is past when any economist should give the slightest countenance to their employment for manipulating trade, or for interfering with the natural tendency of exchange to increase utility."

associated with a larger volume of trade may bring in larger total gain than a smaller volume of trade with a higher rate. On the basis of his assumption of simple production and constant comparative labour costs in the trading countries, Mill could easily calculate the total gain from international trade and divide it between the parties according to the rate of interchange. But when that assumption is dropped, the terms of trade cease to form by themselves an index of the foreign trade position. This is evident from the fact that a higher rate of interchange may be due either to a more intense demand for a country's exports or it may be due to a sharper operation of the law of diminishing returns in regard to the production of the export good. The condition of demand *and supply at home remaining the same, a more intense* foreign demand clearly secures for a country a greater gain. Thus one can easily understand Mill when he says that the greater the foreign demand for a country's product the more favourable is the rate of interchange and the greater is its gain. But when, relying on the terms of trade as the criterion of gain, he adds further that the gain from trade is lower if the home demand for foreign product is greater, and from this deduces that the richest country gains the least form foreign trade, he is certainly putting emphasis at the wrong place.

3. Yet there is a certain degree of truth latent in many of Mill's propositions in these contexts a precise formulation of which is aided by the geometrical apparatus—the demand-and-supply curves—introduced by Marshall.[1] Marshall takes the case of two countries, England and Germany, trading in two commodities, cloth and linen, and constructs the individual demand and supply curves of the trading countries on the basis of diminishing marginal significance

[1] Cf. *Pure Theory of Foreign Trade* (L.S.E. Reprint); also, *Money, Credit and Commerce*, Appendix J.

of one commodity in terms of the other.[1] If, as in the following diagram, cloth is measured along the X axis and linen along the Y axis, and if England is supposed to have a comparative advantage in the production of cloth and Germany in the production of linen, then the willingness of

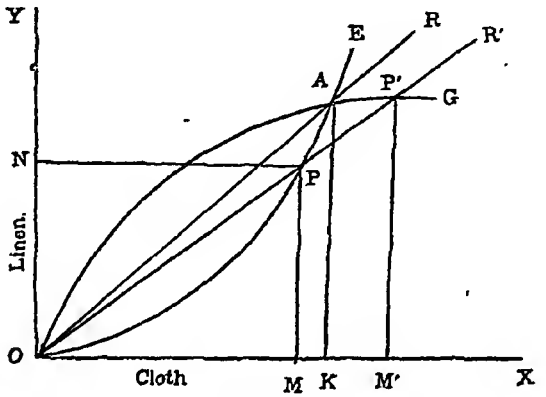

Fig. 6.

England to exchange cloth for linen can be represented on the curve OE and the willingnesss of Germany to exchange linen for cloth can be represented on OG. At any ratio of exchange, tan ∠ R'OX, England is willing to exchange OM quantity of cloth for PM quantity of linen. At this ratio,

[1] In *Pure Theory*, Marshall retains the classical example of two commodities, cloth and linen. In *Money, Credit and Commerce*, he takes the case of *representative bales of E-goods and G-goods*, which stand for a bundle of all goods exported from each country measured in terms of a common unit. A common unit is required because the specific contents of these bales may alter in the process of exchange. Variations in the terms of trade surely tend to alter the range of goods under export and import. Marshall chooses a constant quantity of labour and capital as the unit. This seems again to be an extension of the 'real cost' analysis. When diverse factors enter into production, how are they to be reduced to one common denominator ? In terms of the relative marginal productivity ? But the relative marginal productivity of factors would also vary with variations of income consequent on variations in the terms of trade. It seems no improvement is made by substituting 'factors' for 'goods'.

however, as is indicated by the OG curve, Germany would demand OM' quantity of cloth; and in order to induce England to increase her sales of cloth Germany would have to offer a more favourable rate. Equilibrium is obtained at a point at which the two curves meet, that is to say, at the rate at which the demand-and-supply condition of England equates the demand-and-supply condition of Germany. At the ratio, tan ∠ ROX, England is willing to exchange OK quantity of cloth for AK quantity of linen, and Germany is also willing to offer AK quantity of linen in exchange for OK quantity of cloth. Trade is in equilibrium at the point A.

Now, in what manner precisely are these curves derived, and what clue do they offer in regard to the ascertainment of the net benefit from foreign trade?

Take the case of the curve OE. "N being any point upon OY, let it be determined from a knowledge of the circumstances of England's demand for linen, what is the number of yards of cloth, the expenses of producing and exporting which will be covered annually by the proceeds of the sale in England of an amount of linen represented by ON".[1] If OM be the required amount, then the curve OE will be the locus of P, the point at which the perpendicular through M on OX meets the perpendicular through N on OY. And similarly for the curve OG. This is how Marshall describes his process of deriving the 'willingness-to-trade' curves. It is clear from this description that the curves take into account not only the 'opportunities' that are open in the respective countries for the production of the goods in question but also the relative valuation conferred on them by the consumers. In other words, they take account not only of the 'technical' rate of substitution but also of the 'psychological' rate of substitutaion between the goods in

[1] *Pure Theory*, p.7.

relation to the respective countries.[1] The scope of profitable trade depends upon the rate at which the home cost ratio converges towards the foreign cost ratio and also upon the rate at which the marginal significance of the foreign product falls in relation to that of the home product. The shape of the curves depends upon these two circumstances.

[1] The process of deriving these curves has been shown in a most legant way by A. P. Lerner in his *Diagrammatical Representation of Demand Conditions in International Trade* (*Economica*, Aug. 1934). Taking ne most representative case of increasing costs out of his diagrams the process is shown below (Fig. 7.).

Cloth is measured along the X axis, and linen along the Y axis. ab is the production opportunity curve showing the 'technical' rate of substitution in England between cloth and linen, and the convex lines form the consumption indifference map. If the rate of exchange between cloth and linen is represented by the slope of the line R, which is tangent to ab at Q and also to one of the indifferonce curves at A, then the

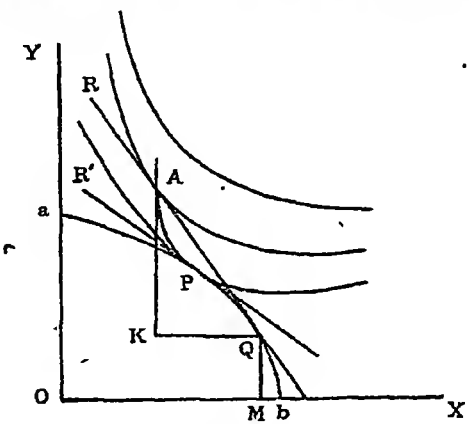

Fig. 7.

locus of A will be the demand-and-supply curve of England of which a section, AP, is shown in the figure. If the rate of exchange is represented by the slope of the line R', which is tangent to the P. O. C. and one of the indifference curves at the same point, P, the country is indifferent as to whether it should enter into foreign trade at all. At the rate, tan ∠ RQK, the country is enabled to move to a higher indifference curve—from P to A—and to exchange KQ quantity of cloth for AK quantity of linen. The

Now, Marshall's demand-and-supply curve in relation to international trade obviously represents only that section of the total offer curve of one good in terms of another which is relevant for foreign trade ; the volume of trade between the goods within the country is left out of account.[1] The relation between this curve and the 'integral demand curve' referred to in the previous chapter is that there the consumer is assumed to start with a given money income and is found to attain a more and more preferred position as the price of the good that he purchases (the objective rate of exchange between the good and money) becomes lower ; whereas here the 'country' is assumed to start with a given amount of resources which can be applied for the production of the goods in question at different rates.[2] If the 'international price' (the objective rate of exchange from the point of view of any particular country) coincides with the equilibrium rate within, the country does not benefit from foreign trade.[3] Starting form the origin where she is indifferent as to whether she should enter into foreign trade, the country, England, for example, attains a more preferred position as she moves higher up along the curve OE. If

diagram may also be used to show which good forms the object of export and which good of import. In this particular case we have assumed that the circumstances are such that England exports cloth and imports linen. But if the condition of foreign demand and supply is such that the rate of exchange is represented by the slope of a straight line which touches ab to the left of P and also one of the higher indifference curves to the right of P (not shown in the diagram), then England will find it profitable to export linen and import cloth.

[1] Translated in terms of the diagram given in the preceding footnote the origin of the OE curve is P.

[2] If one of the goods cannot be produced within the country then its value equivalent in terms of other goods is to be taken into consideration. See for an explanation of the opportunity cost principle, Chap. V, above.

[3] The situation would be analogous to the case of an individual if the price line in the diagram given in the previous chapter (Fig. 4) were assumed to be tangent to the individual's zero indifference curve.

the rate of exchange is tan ∠ R'OX, then England is enabled to move from O to P. As it is, England attains a still more preferred position, A.

4. This is what Marshall tries to emphasize with the aid of his technique of consumer's surplus. In representing the net benefit from foreign trade Marshall, it will be remembered, compares in the case of each country the amount of home produce which 'it would be worth her while to pay' for a given amount of foreign produce and the amount which she does pay.[1] Whether the country is considered as a producing unit with respect to the export good or a consuming unit with respect to the import good, the technique employed in estimating the gain from foreign trade is the technique of consumer's surplus. As it is all a question of *rate*, the gain may be represented as an economy of cost or sacrifice in terms of the export good in securing a given amount of income, or it may be represented as consisting of an extra amount of income secured against a given amount of productive resources. All that the technique of consumer's surplus implies is that at the point, A, England is *better off* than at the point P,—that as she moves higher up on the curve OE she attains a more and more preferred position, considering the internal cost ratio between cloth and linen and the relative demand for them.[2]

[1] See *Money Credit and Commerce*, pp. 161-163; also Appendix J, Secs. I and 3. It should be noted that the calculation in the numerical illustration provided therein contains an arithmetical slip, which was pointed out by Allyn Young. See Marshall on Consumers' Surplus in International Trade, *Quarterly Journal of Economics*, Nov. 1924.

[2] It should be noted that the cost ratio between goods is not absolutely independent of the relative demand. The kind of commodities actually produced influences the tastes of individuals, and the relative valuation placed on goods by consumers is to be taken into account in computing the cost ratio between goods if the latter is not to have too 'technical' a flavour. "The demand of each (country) has its origin in the desires of her people to obtain certain goods from abroad; and her supply has its origin

But what is the significance of the 'line of preference' when it relates not to an individual but to a 'country' which consists of heterogeneous groups whose interests may be divergent? In the case of an individual, the system of indifference curves being given, his line of preference is unambiguous. These curves indicate by their slopes at different points the rate at which he is ready to substitute one good for the other, the different points on each curve denoting by their co-ordinates, combinations of the goods as between which the individual is indifferent in the matter of choice. As he is enabled owing to a favourable market rate to move to a higher indifference curve he gets better off. With a given amount of income he gets a larger amount of satisfaction. In the case of a country, however, movement along the demand-and-supply curve has not that unambiguous significance. In the process of foreign trade while certain groups within a country move up the line of preference, certain others may be moving down. If, for example, there are any factors which are 'specific' to the production of the import good, clearly the extension of foreign trade

in her facilities for producing things which the people of the countries desire. But her demand is, in general, effective in causing trade, only in so far as it is backed by her supply of appropriate goods; and her supply is active, only in so far as she has a demand for foreign goods. Thus the demand of each stimulates the supply of the other; and the demand of each is made effective by its own supply ..... The terms of international trade can properly be said to be governsd by the relations of international demand; but, with equal correctness they can be said to be governed by the relations of international supply. It seems best to speak of them as governed by international demand and supply". (*Money Credit and Commerce*, p. 160). Marshall thus emphasizes the nature of inter-dependence of demand and supply in international trade. This answers Graham's objection that the 'neo-classical' economists put too much emphasis on 'reciprocal demand'. (The Theory of international values, *Quarterly Journal of Economics*, 1932.). It is true, as Graham suggests, that international value is determined by 'linked competition'. But that only means in terms of Marshall curves that these curves are more elastic. See below.

means a loss to them. Again, those people who are largely consumers of the export good may suffer as a consequence of foreign trade. But our criterion of gain being an increase in national income it is a justifiable inference that a country *gains* as circumstances favour a movement upward along its demand-and-supply curve.[1] In the second place, the gain here spoken of cannot be considered in terms of satisfaction, as it could in some sense be in the case of an individual. All that we can say is that the line of preference of the country is reflected in the behaviour of individuals in the market. As such the analysis is necessarily *ex post*.

These considerations should by themselves be sufficient to show that the gain from foreign trade cannot be measured, that it can only be ranked as *greater or less*. If, as we have seen, it is difficult to measure the gain from exchange in the case of an individual, it is absurd to talk of such measure when it is a question of trade between 'countries'.

Nor should the Marshall curves be taken to contain any suggestion to that effect. As we have indicated, Marshall's purpose in this context is not to *measure* the gain from exchange in terms of surplus satisfaction. Although he applies the technique of consumer's surplus in emphasizing the nature of gain derived from foreign trade, yet, as is evident from the whole trend of his analysis in this context, he does not intend to compare the gain from one line of exchange with the gain from another line as he does in his study of domestic trade.[2] Unlike the unit-demand-curve

[1] The losses that the individual members of community may encounter on account of the opening up of foreign trade ultimately raise the problem of distribution of income. Any cause that increases the total income of a community may certainly be made to yield a benefit to all the individual members through an alteration in distribution.

[2] It is noteworthy that Marshall does not use the term 'consumer's surplus' in connection with his study of gain from international trade. It may be that he wanted to avoid confusion between such gains and 'con-

which he uses in his study of 'domestic' trade, and which is drawn with reference to each particular commodity on the assumption that the marginal utility of money is constant and the income elasticity of demand for the good is zero, here in the case of equilibrium of international trade no such assumption is made. In drawing the OE and OG curves it need not be assumed that the different points on the curves are independent of any relation between the goods whose demand they represent and other 'domestic' goods; it is not necessary, as Edgeworth pointed out referring to these curves, that "while the rate of exchange represented by the curves is varied, the rate of exchange between one of the ordinates and all other articles.....remains constant. Rather a movement along a supply and demand curve of international trade should be considered as attended with re-arrangement of internal trade; as the movement of the hand of a clock corresponds to considerable unseen movements of the machinery".[1] Nothing like an 'invariable standard' is invoked in estimating the magnitude of the gain from foreign trade. The technique of consumer's surplus as applied here is just a formal representation of the fact that as reflected

sumer's surplus' as an 'economic measure' of surplus satisfaction. Edgeworth and Allyn Young used the term in this context in their review of Marshall's *Money, Credit* and *Commerce*. See Edgeworth's Review in the Economic Journal, 1923; and Allyn Young, Consumers' Surplus in International Trade, *Quarterly Journal of Economics*, November 1924.

[1] *Papers Relating to Political Economy*, Vol. II, p.32; Cf. also Haberler, *The Theory of International Trade*, p.152. It is, however, curious that Edgeworth while recognising here the interdependence between the terms of foreign trade and the prices of domestic goods invokes in another place the assumption of constant marginal utility of the export good in estimating the net gain from foreign trade. See his Review of Marshall's *Money Credit and Commerce* (Economic Journal, 1923) where he represents the 'surplus' gain from international trade as being equal to the straight line intercepted between the point at which the price line from the origin touches an indifference curve and the point on the zero indifference curve at which the gradient is parallel to the price line.

in the behaviour of a country against variations in the rate of interchange the gain from international trade is *greater* or *less* according as the country is situated at a point on its demand and supply curve which is *farther from* or *nearer to* the origin.[1]

5. These curves offer a clue to an examination of the economic effects of changes in the demand and supply conditions in international trade. Although the net gains secured by any country from foreign trade cannot be measured, whether they are greater or less on account of 'changes in data' can be stated in qualitative terms.

Now, if the shift is in the foreign offer curve the effect is unambiguous. As Edgeworth has shown, a downward shift in the curve due to an impediment means a loss to the home country, and an upward shift means a gain. Given a country's own offer curve, the farther off from the origin the foreign offer curve cuts it, the greater is the gain accruing to it.[2]

When, however, there is a shift in the country's own offer curve the effect is less obvious. It depends upon the degree of elasticity in the foreign curve. When the foreign curve is inelastic a country may improve its position through an impediment to trade,—the case being stronger when the impediment takes the form of a duty on exports or imports.[3] And, by a parity of reasoning, it finds itself worse off in the event of an improvement in home production which shifts its own curve to the right.

[1] If this interpretation of Marshall's theory of gains from international trade is correct, then Allyn Young's criticism of it in the Note referred to in a previous footnote loses much of its force.

[2] Diagrams in proof of this and the following cases are given at the end of this chapter. Cf. also the following Appendix.

[3] Cf. Marshall, *Money, Credit and Commerce*, Bk. III, Chap. VIII, Sec. 2; also Appendix J. Against the loss of consumers is to be weighed the gain of the State in the form of tax receipts.

This is the extent of validity that attaches to Mill's propositions concerning restrictions on foreign trade.[1] There is in fact nothing peculiar in it. If the demand for a particular commodity happens to be inelastic over a certain range, then surely the group of producers engaged in it may increase their total income through a policy of restriction.

But, as Marshall repeatedly emphasized, cases of inelastic demand for a country's product in the international market are extremely rare and are more of theoretical interest. In the first place, the condition of demand in the international market for the product of a country competing with other countries is analogous to the demand curve of a particular firm rather than to the demand curve of an industry as we find in the case of domestic trade. Although the demand for the product as such may be inelastic, from the point of view of any one country it would be more elastic, unless the country happens to have a monopoly with respect to the commodity. Even then,—and this leads us to the second limitation of the case,—any restriction on a particular commodity will have its repercussion on other exports. Hitherto we have considered the case of one export good and one import good. When, however, exports and imports in general are taken into consideration it is most improbable that the offer curves should be inelastic.

It is therefore in exceptional circumstances that a country may derive a benefit at the expense of foreigners. In general

[1] Once the income elasticity of demand is allowed we must also grant the possibility of multiple positions of equilibrium. In such cases obviously a country may secure a better position through interference. Mill, it will be remembered, anticipated this possibility (*Principles*, Bk. III, Chap. XVIII, Sec. 6); but he could not lay down precisely the conditions which might lead to such a situation. For an interpretation of such peculiar cases see Marshall, *Money Credit and Commerce*, Appendix J, Secs. 9-10. The possibility of these cases, although the cases are exceptional, disproves Jevons' unqualified contention referred to earlier, that perfect freedom of trade leads to maximum utility.

the interest of one nation is bound up with the interest of others, and any cause that impairs the total productivity of the trading groups taken together can hardly be beneficial to one group in particular. Fortunately in the generality of cases one need not be an internationalist in order to be a free-trader.

Diagrammatic Notes

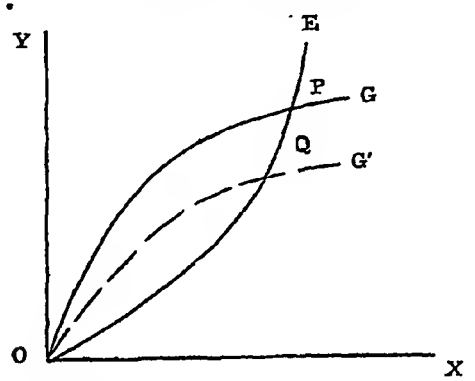

Fig. 8

Fig. 8. For England P is a better position than Q. An impediment shifting OG to OG′ means a diminution of benefit to E. Conversely, an improvement shifting OG′ to OG means an increase of gain. These effects are independent of the shape of the curves.

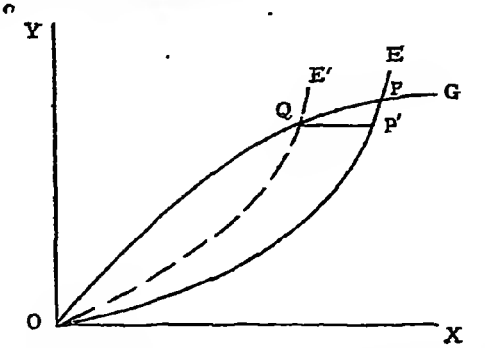

Fig. 9

Fig. 9. This is a case of an impediment shifting the home curve from OE to OE′, the foreign curve being elastic. P is a better position than P′.

Q under the new condition is the same as P'. ∴ Q is inferior to P. The impediment means a diminution of gain. Conversely for an improvement.

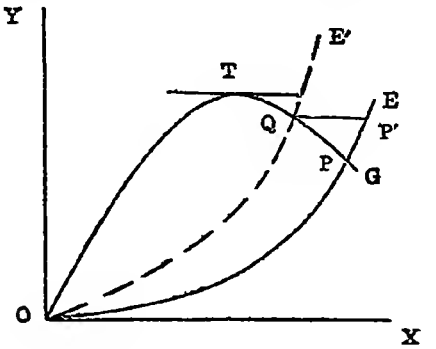

Fig. 10

Fig. 10. The foreign curve being inelastic, an impediment is beneficial to the home country. P is inferior to P' and hence to Q. Conversely for an improvement. The curve OG is inelastic beyond T where the tangent is parallel to the X axis.

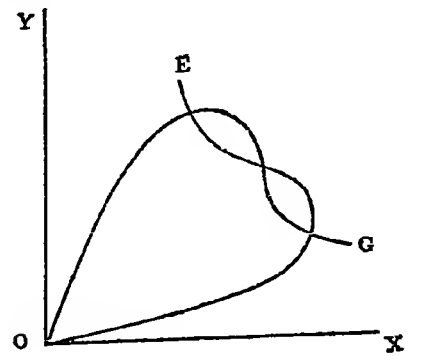

Fig. 11

Fig. 11. This figure demonstrates that if the curves are both inelastic, there may be multiple positions of equilibrium.

APPENDIX

ON THE ELASTICITY OF RECIPROCAL DEMAND AND TERMS OF INTERNATIONAL TRADE

The following note proposes to examine a recent criticism of certain propositions of Marshall concerning the relation of the elasticity of reciprocal demand and terms of international trade, and to offer a general solution of the problem. Marshall, it will be remembered, considers the equilibrium of trade between two countries, E and G, and asks what happens if the demand of one of the countries, e.g., E, for the other country's goods in general increases or decreases. Obviously an increase of demand turns the terms of trade against E, and the opposite happens if there is a decrease of demand. But how far does the movement go? The answer, Marshall says, depends upon the relative elasticities of demand of the two countries for each other's goods. On the assumption that the demand curves of both the countries belong to the 'Normal class',—every point on the curves representing, that is to say, an elasticity greater than unity,—Marshall lays down the following propositions.[1]

Regarding an *increase* of demand,

A. The elasticity of E's demand curve being given, the more elastic the G-curve, the less favourable are the terms of trade to G and the less unfavourable to E; and *vice versa.*

B. The elasticity of G's demand curve being given, the more elastic the E-curve, the larger is the expansion of both exports and imports and the more unfavourable are the terms of trade to E, her exports increasing in a greater ratio than her imports.

[1] Marshall, *Money, Credit and Commerce*, Appendix J, Secs. 4-6; also Chap. VIII.

Regarding a *decrease* of demand,

C. The elasticity of E's demand curve being given, the more elastic the G-curve, the less unfavourable are the terms of trade to G and the less favourable to E; and *vice versa*.

D. The elasticity of G's demand curve being given, the more elastic the E-curve, the larger is the contraction of both exports and imports, and the less favourable are the terms of trade to E, her exports decreasing in a smaller ratio than her imports.

Propositions A and C do not cause any trouble. Whatever may be the nature of the increase or decrease of demand on the part of E for G-goods, in so far as its effect depends upon the elasticity of the foreign curve, the conclusions are unambiguous. The increase of demand, for example, leads to an expansion of the volume of trade. Now, the elasticity of E's demand being given, this expansion will be the larger, the more elastic G's demand is; and the larger the expansion of trade, the higher up on E's new demand curve lies the new point of equilibrium and the less unfavourable are the terms of trade for E and the less favourable for G. Marshall's propositions in this context remain undisputed, and need not, therefore, detain us.

In so far, however, as the effect depends upon the elasticity of E's own demand curve, the conclusions are less straightforward, and they have led to a rather serious misunderstanding. Propositions B and D have recently been challenged. Prof. Viner, following Graham, has declared them to be wrong.[1]

We shall first take the case of an increase of demand. As Viner's presentation of the problem is more precise and hence easier to handle, we shall take his construction for examination although the comments that will be made will apply equally to Graham.

[1] Cf. Graham, "The Theory of International Values" *Quarterly Journal of Economics*, XLVI (1932); and Viner, *Studies in the Theory of International Trade*, pp. 544-45. Viner, it should be noted, does not directly examine Prop. D. But it is not difficult to see his attitude by implication.

Marshall established his propositions with the help of the following diagram.[1]

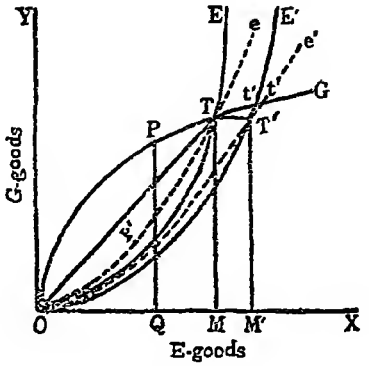

Fig. 12

G-goods are measured along the vertical axis, and E-goods along the horizontal axis.[2] OG and OE are the reciprocal demand curves of G and E respectively, each representing at any given point the total exports the country in question is just willing to offer against a given total of imports. Thus P on OG indicates that G is willing to offer as much as PQ of her own goods against OQ of E-goods, while P′ on OE indicates that E is willing to export OQ of her goods if only she gets P′Q of G-goods. Equilibrium of trade takes place at T where the two curves intersect. In equilibrium OM of E-goods is exchanged for TM of G-goods the terms of trade being TM/OM, or tan ∠TOM. E's demand now increases, and she is willing to offer, let us say, one-sixth more than the previous amount of her goods for any given amount of G-goods. E's demand curve shifts to the right and passes through a point T′, so taken on a horizontal line from T that TT′ is equal to one-sixth of OM. As the curve OG represents, according to hypothesis, an elasticity of demand greater than unity, T′ must lie below OG, as shown in the figure. Now,

[1] For the sake of uniformity I have altered the lettering, where necessary, in this and also in the following diagrams.

[2] The difficulty concerning the notion of 'E-goods' and 'G-goods' is waived here. Cf. footnote on p. 144, above.

whatever may be the shape of the original E-curve, if it passes through T, the corresponding curve representing the same increase of demand must pass through T'. The new equilibrium, therefore, takes place at a point where OG intersects the new E-curve as extended beyond T'. A number of such curves corresponding to different shapes of the original E-curve can be contemplated, and the position of new equilibrium can be shown to depend upon the shape of such curves. t and t' are two alternative equilibrium points, both of which represent terms of trade less favourable to E than TM/OM. Yet at t where OG intersects the relatively more elastic E-curve, the terms of trade are more unfavourable to E, and the volume of both exports and imports larger, than at t' where OG intersects the relatively less elastic E-curve.[1]

Prof. Viner confutes these propositions in terms of the following construction.

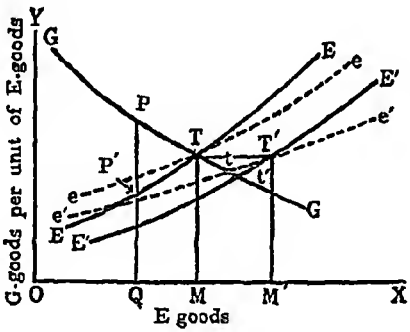

Fig. 13.

On the vertical axis Viner represents the terms of trade (G-goods per unit of E-goods), and on the horizontal axis the quantity of E-goods. GG is G's demand curve showing, for example, that at a price, TM (G-goods) per unit of E-goods, G demands OM of E-goods. As the price falls, G's demand expands. GG is, therefore, downward-sloping. Further as G's demand for E-goods has an elasticity which, according to hypothesis, is greater

[1] As a rule, the flatter the curve the more elastic is the demand that it represents

than unity, the total expenditure of G in terms of her own goods increases as the price of E-goods falls. On the other hand, E's demand (or supply) curve, EE, is upward-sloping, indicating that at higher prices of E-goods—or, in other words, at lower prices of G-goods in terms of E-goods—E is willing to buy more of G-goods, and —the curve being 'normal'—further that a larger quantity of G-goods is demanded against a larger total expenditure in terms of her own goods.

A formal, though not any material, difference between these 'terms of trade' curves, as Viner would call them, and the Marshall curves should be noticed. On OG (Fig. 12), for example, is given the total of G-goods, as measured on the vertical axis, that G is willing to offer for a given total of E-goods measured on the horizontal axis, the ratio between the two amounts being the rate of interchange, or the terms of trade. On GG (Fig. 13), on the other hand, we get the rate of interchange (G-goods per unit of E-goods) and the total of E-goods demanded at each rate straightway; but the total offer of G is given by the rectangle formed by the two co-ordinates at a given point. Similar considerations apply to the E-curve. Thus for an amount OQ of E-goods, G is willing to offer as much as $PQ \times OQ$, while E is just willing to accept $P'Q \times OQ$, of G-goods.

The equilibrium is at T where the two curves intersect. When E's demand increases the E-curve shifts to the right "by a uniform *percentage* at all points on the original curve". E'E', the new curve, according to Viner's construction, is such that if from T on EE a horizontal line is drawn which cuts it at T', then TT'/OM, or MM'/OM, represents the given percentage by which the demand increases. Thus whatever may be the shape of the original curve, if it passes through T, the corresponding curve representing the same increase of demand will pass through T'. The new equilibrium takes place at a point where GG cuts the new demand curve of E. Taking two alternative hypotheses concerning the elasticity of the original E-curve Prof. Viner shows that t, the equilibrium point associated with e'e' which is the more elastic curve, represents a relatively less unfavourable rate of interchange to E than t' which is associated with the

relatively less elastic curve, E'E'. Since the curve GG is 'normal' it can also be easily deduced that a larger volume of exports and imports is associated with a less elastic than with a more elastic E-curve.

All this is surely opposed to Marshall's findings. Yet that does not mean that there is anything wrong in Marshall's construction. Prof. Viner's own findings are all right so far as they go. But he is seriously mistaken in thinking that they also prove that Marshall is wrong. The charge that he brings against Marshall's diagram in this connection[1] is also entirely unwarranted.

[1] "The unnecessary complexity of Marshall's diagram seems to have concealed from him the fact that it provides no answers to the questions which he was putting, for the diagram.....shows three *original* English curves, different in locus as well as elasticity, and fails to present a comparison of the effects of an *increase* in an original English curve according as that original curve has high or low elasticity." (*Op. cit.*, pp. 545-46).

It is difficult to see what Prof. Viner means by all this. Marshall's lettering in Fig. 12 (*Money, Credit and Commerce*, p. 343) has been a little unfortunate. The curve aE has just a chance of being confused with the original curve OE in Fig. 11. May be it is this that has misled Prof. Viner into thinking that aE, aE', and aE'' in Fig. 12 are continuations of three original English curves. But he should have noticed that these curves pass through a (corresponding to our T'), and thus represent E's demand under new conditions.

Personally I feel convinced that the curves which Marshall introduces in this context,—Integral Demand Curves, as they are called—are neither unnecessarily complex, nor do they conceal any matter which Prof. Viner might want to bring out. The analysis contained in the text will be itself sufficient evidence, I hope, of the potency of these curves. On the other hand, even if Prof. Viner should feel that they are complex—although that is a matter of opinion—he ought surely to concede that for Marshall's own purposes they were a necessary apparatus. Apart from the question of symmetry to which Prof. Viner alludes but does not attach as much importance as Marshall would, there is one rather more important point to be taken into account. As is well known, the so-called Unit Demand Curves that Marshall employs in his study of 'domestic values' are frankly based on the assumption that one of the goods, namely, money—in terms of which all prices are reckoned—has consitant marginal utility. It is thus significant that he should have kept them distinct from the type of curves employed in the study of 'foreign trade', where apparently no such

For, Prof. Viner is unaware of the fact that his findings are based on an assumption concerning the nature of E's increase of demand which is different from Marshall's assumption. It is true that both refer to an increase in E's offer by a uniform percentage. But, whereas in Viner's construction the increase is shown against a given *price per unit*, in Marshall's construction it is considered in relation to a given *volume of imports*. TT' in Viner's diagram (Fig. 13) measures the increase in E's offer with reference to the old equilibrium price, TM,—whereas in Marshall's diagram (Fig. 12) it measures the increase with reference to the old equilibrium imports at T, the price obtaining there, namely, T'M'/OM', being less than the old equilibrium price, TM/OM.

It is surprising that Prof. Viner should have failed to notice this. If he had correctly represented Marshall's position he would have observed that his own curves also yielded results corroborating just those propositions that he thought he disproved. The Marshall case as represented on the Viner diagram would stand as follows.

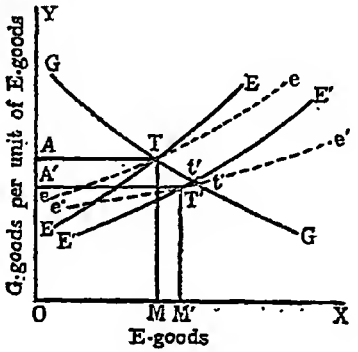

Fig. 14;

GG is again the demand curve of G, and EE the relatively less elastic demand curve of E. TM is the equilibrium price per unit of E-goods in terms of G-goods; OM, the total volume of

assumption was made. For Marshall's own justification for the use of these distinct types of diagram, see his *Pure Theory of Domestic Values* (London School Reprint, No. 1), pp. 1-2.

exports, and AOMT, the total volume of imports of E. Now E's demand increases in such wise that E is willing to offer one-sixth more of her goods against a given volume of imports in terms of G-goods. To get the point which will represent this increase of demand with reference to T, take a length $MM' = \frac{1}{6}$ OM, and draw MT' perpendicular upon the horizontal axis, such that A'T'M'O = ATMO, or $M'T' = \frac{6}{7}$ MT.[1] T' is then the required point. The new curve representing the given increase of demand will pass through T'. For a given volume of imports, A'T'M'O (=ATMO) of G-goods E is now willing to offer $OM' (= \frac{7}{6}$ OM) of her own goods. And this is precisely the hypothesis of Marshall. Thus any E-curve passing through T, the original equilibrium point, when shifted in view of the given increase of demand, will pass through T'.

Now, it should be noted—and this is vital for the validity of Marshall's propositions—that T' must lie below the G-curve. According to hypothesis, GG represents an elasticity of demand which is greater than unity. The total expenditure of G at lower prices of E-goods is, therefore, larger than at higher prices. This means that the rectangle formed by the co-ordinates of any point on GG to the right of T must be greater than ATMO. But T' which lies to the right of T has co-ordinates forming a rectangle which, as we have seen, is equal to ATMO. Therefore T' which represents the given increase of E's demand with reference to the old equilibrium point, T, must lie below GG. The rest is simple. As before, have a relatively more elastic curve, ee, passing through T. E'E' and e'e', the new curves corresponding to EE and ee respectively, will both pass through T'. t', the new equilibrium point on the relatively less elastic demand curve, represents a relatively less unfavourable rate of interchange for E, and a smaller volume of exports and imports; while t, the new equilibrium point on the relatively more elastic curve, represents a more unfavourable rate of interchange and a larger volume of exports and imports. All this conforms to Marshall's findings.

[1] $OM' \times M'T' = OM \times MT$. Now, $OM' = OM + \frac{1}{6}OM = \frac{7}{6}$ OM. $\therefore \frac{7}{6}$ OM $\times$ $M'T' = OM \times MT$, or $M'T' = \frac{6}{7}$ MT.

On the other hand, the Viner-Graham propositions, as tested on the Marshall diagram, do stand when the precise assumption on which they are based is adhered to. The assumption, let it be repeated, is that for a given price, E's offer increases in a given percentage.

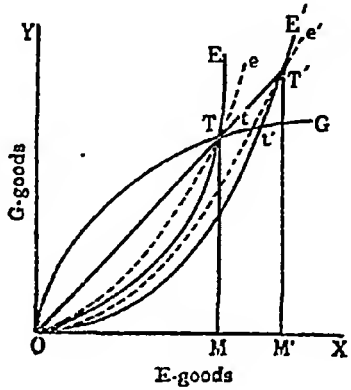

Fig. 15.

In the above figure T is as usual the old equilibrium point, and T' (on OT produced) represents the increase of demand such that, T'M' being drawn perpendicular upon OM produced, MM'/OM=k, where k is the given percentage. The new curve thus passes through T'. t is the new equilibrium point associated with a relatively more elastic demand curve and represents a relatively less unfavourable rate of interchange and a smaller volume of exports and imports, while t' which is on the less elastic demand curve represents a more unfavourable rate of interchange and a larger volume of exports and imports.

The demonstrations given above should make it clear that it is all a question of whether T' which is taken as the standard point representing the increase of demand with reference to the old equilibrium point, T, lies above or below the G-curve. The Marshall case implies that it should be below, while the Viner case shifts it above the G-curve. Hence the difference in their findings. To put the matter in plain language, it is a question of the initial impact of the increase of E's demand. If the initial

impact is such an increase in the volume of exports (and imports) that the establishment of the new final equilibrium involves a relative contraction thereof, the more responsive (or elastic) E's demand condition is, the less unfavourable will be the final rate of interchange to her. On the other hand, if the initial impact is such that the process towards the new equilibrium involves a further expansion of the volume of exports (and imports), the less responsive E's demand condition is, the less unfavourable will be the final rate of interchange to her. And all this depends upon the nature of increase of demand one has in view. Marshall analyses the latter situation, while Viner, without being aware that he is making a different assumption, refers to the former situation. Viner is thus surely right in his own way; but he is wrong in thinking—as he does—that Marshall is wrong.

The above methods can now be easily extended to an examination of the effect of a decrease of demand. Viner does not examine this aspect directly. Graham does, and reaches conclusions which are again opposed to Marshall's. Here again the difference can be shown to arise from a difference in assumption concerning the nature of decrease of demand.

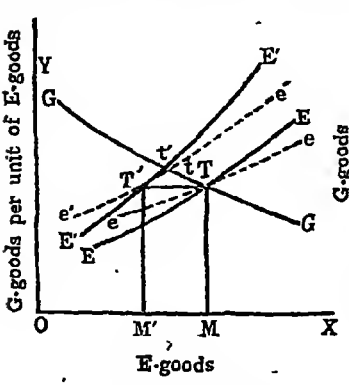

Fig. 16.

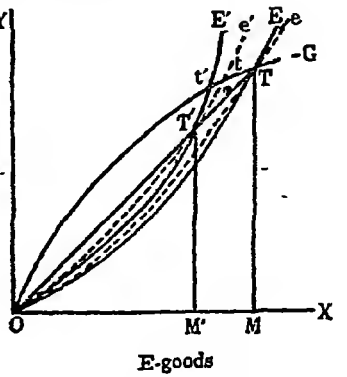

Fig. 17.

E's demand decreases. The terms of trade move in favour of E. In Figs. 16 and Fig. 17 is represented the case where the volume of exports and imports is relatively larger and the terms

of trade more favourable to E when the demand for G-goods is more elastic than when it is less elastic. T is the old equilibrium point and T' represents the decrease of demand with reference to T, such that for a given price MT (Fig. 16), or tan ∠TOM (Fig. 17), E's offer decreases by MM', this MM' being in a given ratio to OM. At t, the new equilibrium point associated with a more elastic demand curve, the volume of trade is relatively larger and the terms of trade more favourable to E than at t'. All this is in accordance with Graham's findings.

But the other possibility is there, too.

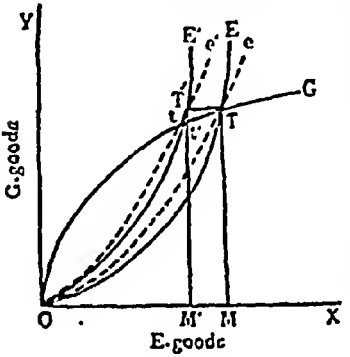

Fig. 18.

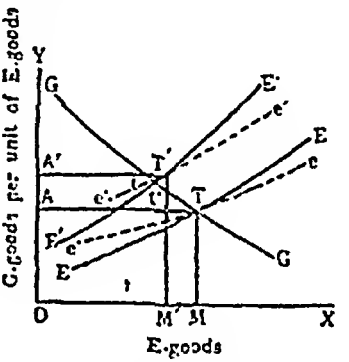

Fig. 19.

In Figs. 18 and 19 is represented the Marshall case where the opposite results follow. For a given volume of imports E's offer decreases by a certain given percentage, say, one-sixth. The construction of Fig. 18 is simple. From T draw a horizontal line TT' towards the left of the original E-curve, such that TT'= MM'=$\frac{1}{6}$OM. In Fig. 19, T' represents the same decrease of demand with reference to T, so that MM'=$\frac{1}{6}$OM, and OM'×M'T' =OM×MT,—from which it follows that M'T'=$\frac{5}{6}$MT. In the present case, T' will lie above the G-curve. G's elasticity of demand being greater than unity, the rectangle formed by the co-ordinates of any point on GG to the left of T is smaller than OM×MT. But according to hypothesis, the co-ordinates of T' form a rectangle which is equal to OM×MT.

Both the diagrams yield the same result, as they should. At t, the new equilibrium point associated with a relatively

more elastic demand curve, the volume of trade is smaller and the terms of trade less favourable to E than at t′ which is the new equilibrium point associated with a relatively less elastic demand curve.

Here again it is all a question of whether T′ is below or above the G-curve. In other words, it is a question of the initial impact of the decrease of E's demand. If it is such that the process towards the new final equilibrium involves a relative expansion of the volume of trade, Graham's findings are justified. If, on the other hand, it is such that a further contraction of the volume of trade is involved in that process, then Marshall's propositions are found to be perfectly all right. And, of course, Marshall's assumption definitely implies the latter phenomenon.

One feels inclined further to suggest that not only is Marshall not wrong, in the contexts which are particularly relevant to the study of the problems in question it is Marshall's construction and the propositions that go with it that are based on a more plausible assumption. If the decrease of a country's demand for foreign goods arises from the introduction of obstacles in the form, for example, of *ad valorem* import or export duties, and if an increase of demand arises from the removal of such obstacles—and indeed Marshall had such things in view while he enunciated his propositions—it is surely more pertinent to consider the percentage change in amounts offered with reference to a previous *total* of imports. But that, of course, is a question of the facts of a particular situation and no more than a passing reference is justified here.

DISTRIBUTIONAL SURPLUS
DYNAMIC ASPECT

CHAPTER IX

FRICTIONAL SURPLUS

1. So far our discussion has been carried on in terms of a static economy. It has been assumed throughout that there are forces in the economic system which bring about a condition of equilibrium between demand and supply,—a condition in which the relative prices of goods are adjusted to their scarcity in relation to demand, and productive factors are distributed among different lines of demand in such manner that these prices correspond to the relative costs which the employment of factors gives rise to. In the sphere of exchange the consumer derives the maximum of utility out of his income by so distributing them that the relative marginal utilities of the different goods purchased are equated to their relative prices. In the sphere of production a similar maximum principle operates through an adjustment of productive resources as between different lines such that their marginal productivities in value terms are equalised. In view of this striving of economic subjects towards maximum satisfaction (or minimum sacrifice) a situation is envisaged in which the total demand equates the total supply in each line of economic activity ; and, as the demand for one commodity derives in the ultimate analysis from the supply of another commodity, this equation means that no incentive is left for any shift in resources from one line to another.

This is a picture of general static equilibrium in which all the elements in the economic system are at rest,—or more properly, in 'steady motion'. Here we fix our eyes on a point at which, given the initial data concerning factor supplies, relative demand conditions of goods etc., the

economic system finds itself in a state of equilibrium ; and we assume further that in the absence of any change in data the equilibrium continues intact through time.

Now, economic equilibrium itself relates to a process through time ; and by fixing upon a point at which an adjustment of economic phenomena has taken place, and by representing the economic process as a uniform and steady flow of wants and activities, we are assuming away the important role of time in the process of that adjustment which we hold in view,—the resistances that are encountered by economic factors in the path towards equilibrium. It is not our purpose here to probe the network of complications that attach to this problem of the 'path'. We shall only refer to certain outstanding features connected with it which are relevant to our immediate problem of economic surplus.

2. Once it is recognised that the economic mechanism does not bring forth ready and instantaneous adjustment of economic factors,—that it takes time of varying length for these factors to transmit their effects into the system, two things stand out as deserving attention,—the problem of anticipation, and the psychological and technical resistances that the economic subjects have to face in their endeavour to adapt their activities to the phenomena anticipated and further to bring these anticipations into correspondence with what actually would occur. Economic activities are directed to the satisfaction of human wants. But just that want which is the object of a particular activity is remote,—often far remote,—from the time at which it is undertaken. Hence the role of anticipation. Cost is undergone with a view to an anticipated value. Investment is made with a view to an anticipated income. Payments are made to hired factors with a view to their anticipated marginal (value) product. Everywhere supply is adjusted not to the

actual demand but to a demand which is anticipated. And as men have no perfect foresight there is always present in any undertaking some degree of uncertainty which creates a divergence between the anticipated and the realised result; —a divergence which again causes shifts in the economic system. Uncertainty is of the essence of economic dynamics. One of the driving forces in the economic system is therefore that of people gathering knowledge with a view to making correct anticipations of the course of events in the market and adapting plans accordingly.

It is thus clear that the extent of divergence between the anticipated and the realised result in any undertaking constitutes a surplus. We may call this Dynamic Surplus. It is the result of change and is itself a potent influence causing change.

Dynamic surplus as thus defined belongs more properly in the field of Distribution, denoting an excess over 'producer's costs'. The tendency of relative costs to correspond to the relative values of goods is a long-run tendency and holds good under a static hypothesis. Over any particular planning period, however, costs are adjusted to the anticipated value of the wares turned out. In a changing environment and in the absence of perfect foresight the actual value of the wares at the time of disposal may diverge from the value with a view to which those costs had been incurred. From the point of view of any particular producer, therefore, this failure of correspondence between the actual and the anticipated value of his wares gives rise to a surplus, —positive or negative, as the case may be. Positive surplus in any undertaking arises on account of an initial pessimism on the part of the producers in that line resulting in a smaller investment of resources than is warranted by the condition of demand. Negative surplus arises on account of an initial optimism resulting in an over-investment. But as soon

as this disharmony between demand and supply is felt there is a tendency towards a shift of resources. The surplus emerging during this period of mal-adjustment is called Net Profit in economic science. It is the result of change accompanied by uncertainty. At the same time its occurrence is an indication of further changes in the economic system. The implication of static hypothesis is that, abstracting from extraneous influences and concentrating on what are called 'endogenous' repercussions of a given system of economic forces, it is presumed that buyers and sellers of goods work through *recontract over time* towards a point at which this surplus is wiped out.

It is often argued that consumer's surplus also belongs to this category,—that like pure profit it is an offshoot of economic change and tends to vanish as things settle down.[1] The idea is that it takes time for a consumer to get adjusted to changing price ratios in the market. When this adjustment takes place there cannot be any consumer's surplus; for just so much of a good is consumed that the total price that is required to be paid coincides with the total utility that is derived from it. Only under dynamic conditions, when consumption is not adjusted to the price ratio in the market, can there arise a consumer's surplus. If there is a sudden fall in the price of a good, the consumer secures a surplus of satisfaction, but he soon gets accustomed to the

[1] Cf. Harry E. Miller, Utility Curves, Total Utility and Consumer's Surplus, *Quarterly Journal of Economics*, Feb. 1927. "The principle of diminishing utility.... operates in such a manner as to differentiate the several 'doses' of a good only during periods of change from long established rates of consumption. Graphically, this means that, except during such periods,... total utility is indicated by a rectangle". Hence, it is argued, "it is only during transition periods ... while a new rate of consumption is bringing the purchasing power which money has for that good into harmony with the individual's marginal utility of money in the general field of exchange—that a consumer's surplus of the sort commonly described would seem to appear".

new price and adjusts his consumption in such manner that the surplus fades away. His long-run expenditure on the good is such that it swallows up the entire utility derived from it.

'But if the analysis of the preceding chapters is correct, then it follows that this dynamic theory of consumer's surplus is a clear misapprehension. Not only does consumer's surplus persist under static conditions, it realises itself more fully in the long than in the short period.

This requires a little more careful examination.

The indifference system of the consumer with reference to which his demand curve is derived is related to the time that is allowed for the adjustment of his preferences. The consumer's short-period preference schedule is different from his long-period preference schedule. For it takes time for the interdependence of the utility of goods to show itself in the consumer's distribution of income to the optimum level. Thus in the example of tea referred to earlier, the demand curve of the consumer representing the comparative utility of tea and money would be different in relation to a short period of imperfect adjustment from what it is under static conditions in which an optimum distribution of money among other goods is assumed. If a sudden price variation in favour of the consumer leaves him better off and affords him a surplus, this surplus brings in an appropriate utility only as time is allowed for a fuller employment of it through necessary re-adjustments in the entire scheme of consumption.

This short period mal-adjustment may arise on account of an inherent obstacle in the adaptation of the uses of goods to changing circumstances. The degree of satisfaction that is derived from an increment of a good depends on how far the consumer is capable of re-shuffling his expenditure on other goods according to the new situation; and this is a

question of time.[1] Given time and the possibility of a readjustment of his expenditure the consumer may look forward to a better position than he immediately secures. A delay in adjustment may also be due to a lack of knowledge of relevant data on the part of the economic subject. Here again if one fails to make a ready adjustment of his consumption to changing price ratios in the market, this mal-adjustment far from being the cause of the emergence of a surplus utility does, on the other hand, stand in the way of his reaping the fullest benefit which the possession of a given amount of resources warrants,—of the ultimate surplus derivable from exchange realising itself fully. It is through experience that the consumer forms his ultimate scale of preference between goods ; and a static demand curve with reference to which a rational distribution of income resulting in a maximum of surplus utility is envisaged assumes a perfect knowledge on the part of the consumer about the relative significance of the goods concerned.

Consumer's surplus is thus a phenomenon connected with economic statics rather than with economic dynamics. Its appearance does not mark a transitional stage in the economic process. On the other hand, it is on a long period static hypothesis that the consumer is seen to be able to make the most of a given situation and to derive the maximum surplus utility from exchange.

3. In discussing the problem of uncertainty and dynamic surplus, therefore, we should turn to the relation between value and producer's costs. It is, as we have seen, a phenomenon resulting from a divergence between the anticipated value and the actual value of a good, or rather,

[1] The distinction drawn by Austrian economists between 'Unmittelbar abhängige Nutzen' and 'Mittelbar abhängige Nutzen' seems to point to this phenomenon. Cf. Rosenstein-Rodan, "Grenznutzen" in *Handwörterbuch der Staatwissenschaften*, Bd. 4.

between the expenses incurred by the producer of a good and the price at which it is sold.

Explicit mention should be made here of the distinction that has figured in economic analysis between Friction and Pure Dynamics. In the one, uncertainty is regarded as a transitional phase which merely delays the attainment of —and yet ultimately leads through stages to—a definite and determinate equilibrium. In the other, uncertainty and its correlate, anticipation, are admitted to be an active element, changing the course of the economic system away from the 'norm' postulated in static analysis. The latter enquiry is more interesting and has come to hold the field in recent years. The significance of this 'pure' dynamic surplus which is peculiarly associated with money economy will be examined in the next chapter. A brief reference may in the meantime be made to the earlier doctrines of Profit as a surplus which were in fact set out essentially against a static background and which we may designate as Frictional Surplus.

Now, in a system of private enterprise as we have today, a mal-adjustment of economic activities may be looked at from two points of view. It may be looked at from the point of view of the relative position of different industries and the distribution of resources as between them. And further, it may be looked at from the point of view of the relative position of the 'entrepreneur' who takes the initiative in business undertakings and the factors receiving contractual payments. In the first case it is a question of over-investment or under-investment of resources in particular lines of undertaking. In the other case it is a question of a relative excess or deficiency in the supply of agents ready to bear uncertainty.

The former type of short period mal-adjustment finds recognition in a letter which Turgot wrote to Hume as

early as 1767.[1] Turgot laid down rather in clear terms the importance of the time element involved in the passage of the economic system from one equilibrium position to another. In the process towards the 'equilibrium of .... values' there are 'frictions' giving rise to fluctuations in 'current' values. The idea is also implicit in the classical theory of market price, as distinct from natural price. The classical economists were mainly concerned with the study of natural price,—the end-position of the economic system, as they viewed it. But they recognised that the attainment of this position is an ultimate tendency in the course of which there might be mal-adjustments giving rise to a divergence between the actual price and the natural price. Clearly, an excess of the actual price over the natural price means a positive surplus, and a deficiency means a negative surplus ; and all this is due to the under-production or over-production of goods.

In all this, however, 'uncertainty' fails to achieve its proper status. The 'variations' that these studies suggest are of a simple type denoting stages in the so-called 'equilibrating process'. We have already given some account of the significance of this type of surplus in our study of Marshall's theory of Quasi-Rent.

One point, however, needs to be mentioned in this connection. In so far as a surplus of this type arises on account of a mal-adjustment of resources as between different industries, it attaches during any period to particular industries rather than to the system as a whole, *so long as the rule obtains that the demand for one commodity derives from the supply of another commodity.*[2] For, if it is a question of the

[1] *Reflections*, (ed. Ashley) pp. 100-109.

[2] It should be noted that this is the assumption on which classical economic analysis was based. As it is a question of the employment of *given* resources, and of *relative* prices of goods, supply and demand are two aspects of the same thing. This doctrine of the coextensiveness of demand

distribution of *given* resources between lines which are *competitive* in respect of demand for those resources, a positive surplus in one line is balanced by a negative surplus in another line. Over any particular economic period investment are made by entrepreneurs in different directions on the basis of certain anticipations of demand for the respective products. But if of these products the demand for one group comes from the supply of another, a wrong anticipation with respect to one leading to profits necessarily implies a wrong anticipation with respect to the other in the opposite direction leading to losses. The actual situation with respect to one group at the end of the economic period is a factor influencing the realisation of anticipation with respect to the other group. Profits and losses are themselves relative concepts in this context.[1] When products are considered in value terms, it is clear that if the actual yield in one line of investment exceeds the prospective yield on which it was undertaken, it means that with respect to products in terms of which its value is reckoned, the actual yield falls short of the

and supply, *i.e.* of production and consumption, was first formulated by J. B. Say. Cf. *A Treatise on Political Economy* (4th edition, Eng. Tr.), Bk. I, Chap. XV. "It is production which opens a demand for products", —so that "... precisely at the same time that one commodity makes a loss, another commodity is making excessive profit." James Mill also demonstrated it in a pamphlet, *Commerce Defended*, 1808. A more clear exposition is to be found in his *Elements of Political Economy*, Chap. IV, Sec. iii. See also J. S. Mill, *Principles* (ed. Ashley), Bk. III, Chap. XIV. The limitation of the doctrine, arising as we proceed to translate relative prices into absolute monetary prices, will be pointed out in the next chapter.

[1] "... if any given price yields a profit, resources will be diverted *to*, and if a loss, *from* the production of that good; the real meaning of profit is simply that resources being used to produce other goods (and valued in other uses) will yield more in the production of the good in question; while similarly, loss means that resources producing the good in question are worth more in other uses (their value being determined by that of the best use)." Knight, *Risk Uncertainty and Profit*, Chap. III, p. 91. In a sense, however, any mal-distribution of resources of this type results in a social loss, for it secures less than maximum satisfaction of consumers' wants.

prospective yield. But if this happens in any particular period, then there is a tendency towards a redistribution of resources. From time to time entrepreneurs gather experiences about the course of affairs in the market, and starting from any given point of time, it is expected that in the absence of any 'exogenous' forces frustrating the anticipations of entrepreneurs, the economic system would work through towards a point of equilibrium at which no profits or losses remain. Profits or losses emerging at the end of any economic period are given data for the next period. The entrepreneurs take stock of these situations in revising their plans for future production. For a time the 'quasi-rents' from certain investments may be higher than what they would be under equilibrium conditions; but at the same time the earnings from some other investments would be lower. But as soon as such a situation occurs there is a shift of resources, and given time the discrepancies are smoothed out. It is a zig-zag course that the system follows in its progress towards equilibrium, and at each point on the course, which represents a planning period, new knowledge is acquired which regulates the action of entrepreneurs. The elimination of the element of uncertainty in static analysis signifies this, that if people have perfect foresight and are possessed of all relevant data so that the equilibrium rate of output of different industries is struck just at the starting point of economising and is maintained through time, then the economic system takes a smooth and regular course, and profits and losses do not arise.[1]

4. Now, although under dynamic conditions, owing to mal-distribution of resources, profits and losses attach to particular industries, there is nothing in what has been said which explains 'profit' as a share in the distribution of

[1] Cf. Marshall's 'true equilibrium rate' in Barter. *Principles* (7th edition), Appendix F, p. 791.

national income. For this we have to turn to the relative position of entrepreneurs as a class and hired factors who receive contractual payments. The entrepreneurs who are the hiring factors undertake to assume the risks that attach to business undertakings and may secure an income which exceeds that which they might secure in the position of hired factors. If they do they secure a profit and this profit may be described as a general category of surplus income. The payments that are made by entrepreneurs to the hired factors are based on what is expected to be their marginal productivity. If these expectations prove at the time of the disposal of the products to have been too optimistic the entrepreneurs suffer loss. If, on the other hand, the expectations prove to have been too pessimistic they secure profit. In either case the entrepreneurs find at the end of the economic period that their expectations have gone wrong,—that they secure out of the proceeds of their business an income which is different from what they had expected,—or in other words, an income which is different from what they would normally earn as hired factors.

This essentially dynamic character of profit has been recognised only in recent years, and the recognition has been the result largely of the introduction of mathematical notions into economic science.

It is well-known that the classical economists identified profits with the earnings of capital, or 'stock', as they used to call it. The element of risk in business undertaking was no doubt recognised. But the unique character of this element,—the role of uncertainty in business undertaking, was not shown,[1] and any non-contractual income was desig-

[1] Here again Cantillon was more explicit than Adam Smith and his immediate followers. Cantillon divides all the inhabitants of a State except the Prince and the Proprietors of land into two classes : Entrepreneurs and Hired People. It is interesting to note that the term 'Entre-

nated as profit. In so far as risk was inherent in business—in the sense of loss of capital,—the owners of capital would normally earn something more than ordinary interest; but this element was not differentiated and the significance of normal conditions was not realised.

With some, Senior and Cairnes, for example, risk forms an element of 'real cost'. Risk, they argue, adds to the 'irksomeness' of business and raises the 'real cost' of production. Remuneration is also proportioned, other things remaining the same, to the degree of risk that labour or capital is exposed to. J. S. Mill analyses profits into three elements,—wages of management, interest, and a payment for risk,—all these being included in the natural price of commodities. Even in Marshall this loose use of the term persists. In fact he goes at pains to prove that normal profit (including a remuneration for risk) is an element of cost of production. Marshall's attitude in this connection is to be explained partly by the fact that he was always eager in appropriating a term to keep close to business usage,—and partly—and this is more important—by the fact that he had an aversion to rigidly static analysis and would insist on discussing "representative" conditions as opposed to limiting tendencies.'[1]

The unique character of profit as a dynamic surplus has been brought out by J. B. Clark. Although Clark does not

preneur' which has been so much in vogue in economic literature was first used by Cantillon. All the entrepreneurs, he says, "are as it were on unfixed wages and the others on wages fixed so long as they receive them though their functions and ranks may be very unequal. The General who has his pay, the Courtier his pension and the Domestic servant who has wages all fall into this last class. All the rest are Undertakers [Entrepreneurs], whether they set up with a capital to conduct their enterprise, or are Undertakers of their own labour without capital, and they may be regarded as living at uncertainty." *Essai* (ed. Higgs), p. 55. See in this connection the whole of Chapter XIII.

[1] See his correspondence with J. B. Clark in *Memorials of Alfred Marshall*, (ed. Pigou).

develop an independent theory of profit, he explains nevertheless how profits cannot arise under static conditions and how it is dynamic conditions that give rise to a surplus above cost. According to Clark, so long as those 'five generic changes'[1] that constitute a dynamic economy do not occur, prices just cover costs and there cannot emerge any profit. In static conditions, under the pressure of competition, hired factors are paid according to marginal productivity. These payments (including payments to the entrepreneur's own resources) exhaust the total product and no residue remains. "Entrepreneurs' profit and residual income are synonymous terms." Static conditions preclude the existence of such profits. Profits may arise from 'frictions', —'variations in the static standards'. But above all they arise from 'changes in the standards themselves'. And he illustrates one of those changes by 'invention'. "Thus an invention makes it possible to produce something more cheaply. It first gives a profit to entrepreneurs and then ... adds something to wages and interest .... When the full fruits of this invention shall have diffused themselves throughout society, the earnings of labour will equal the new standard rate." Inasmuch as there is a 'static standard' even in a dynamic society, profit that arises from change is 'an elusive sum, which entrepreneurs grasp but cannot hold.' It "slips in time through their fingers and bestows itself on all members of society."[2]

Development on Clarkian lines of the conception of profit as a dynamic surplus has been further carried by Prof. Schumpeter. Schumpeter shows—as does Clark—that under static conditions,—under a regular 'circular flow of economic life'—the productive process requires no co-ordination. The function of the entrepreneur is the 'carrying out of new

[1] See *Distribution of Wealth*, Chap. V, p. 56.

[2] *Op. cit.*, p. 405.

combinations'; and this is a dynamic process. Entrepreneurs effect "changes in the productive process in the widest sense, the aim of which is to produce a unit of product with less expense and thus to create a discrepancy between their existing price and their new costs."[1]

But, as Prof. Knight has pointed out, it is not change as such, but uncertainty about this change that gives rise to a residual income of entrepreneurs.[2] Clark and Schumpeter fail to take note of Uncertainty as the essentially dynamic element in the economic system. Take the case of innovations. If the innovator foresees the consequences of his action upon the cost and price of the commodity in question, then what he gets is simply a higher value imputed to his superior capacity. If, therefore, monopoly revenue is outside this category—and so it is, for we are concerned here with dynamic surplus as such—it is difficult to see how the income of the innovator, when he foresees the effect of his innovation, can be called net profit.[3]

Not only do they ignore the role of uncertainty in the evolution of profit, both Clark and Schumpeter definitely reject the so-called risk-theory of profit. In his criticism of Hawley who had been developing contemporaneously a risk-theory of profit Clark argues that if risk-taking gives rise to a special category of income it goes to the capitalist and not to the entrepreneur as such. "It goes without saying that the hazard of business falls on the capitalist. The entrepreneur, as such, is empty-handed. No man can carry risk who has

[1] *The Theory of Economic Development*, p. 133.

[2] "It is not dynamic change, nor any change, as such, which causes profit, but the divergence of actual conditions from those which have been expected and on the basis of which business arrangements have been made." *Risk Uncertainty and Profit*, p. 38.

[3] Foresight about the effect of change is explicitly assumed by Schumpeter. After stating three conditions which create profit, he says, "they are obviously not always fulfilled, and when they are not, and the fact is foreseen, the new business is not organised .... If the conditions are fulfilled, however, the surplus realised is *ipso facto* a net profit." *Op. cit.*, p. 131.

nothing to lose."[1] The obvious criticism is that to the extent that the uncertainty is borne by the capitalist he *is* the entrepreneur. The ultimate responsibility of making decisions devolves upon the person who faces the uncertainty of business. One has here to make a distinction between *taking* the initiative and *being made to take* it. Hawley makes this quite clear when he says that an entrepreneur is he who ultimately bears the consequences of decisions taken in business, and to the extent that an entrepreneur decides to lessen his risks he abdicates his peculiar entrepreneurial function.[2] Further, it should be remembered that an entrepreneur cannot work 'empty-handed'. If he wants to borrow money at stipulated interest he must own resources to protect the lenders, and the the first charge of any possible loss is upon those resources. It is to the extent that these resources do not cover the risk that the entrepreneurial function is shifted to the lenders.[3]

[1] Insurance and Business Profit, *Quarterly Journal of Economics*, Vol. VII, p. 46. Cf. also Schumpeter, *op. cit.*, p. 137. "The entrepreneur is never the risk bearer.... The one who gives credit comes to grief if the undertaking fails."

[2] See Hawley, The Risk Theory of Profit, *Quarterly Journal of Economics*, Vol. VII, p. 464; Enterprise and Profit, *Quarterly Journal of Economics*, Vol. XV, pp. 75-105. For an earlier statement of his theory, see Profits and the Residual Theory, *Q. J. E.*, Vol. IV, pp. 387-396. Prof. Knight gives an excellent summary of the controversy between Clark and Hawley on this question. (*Op. cit.*, Chap. II). But it seems he does a little less than justice to Hawley. It is true that Hawley did not develop the part played by uncertainty in the economic system. But he certainly saw the nature of uncertainty, as the following quotation would show: "Uncertainty and risk are connotations. The one cannot exist without the other. The moment any element is determined, uncertainty ceases. It is therefore only in the undetermined residue that uncertainty inheres. A residue, until finally determined, connotes a risk; and a risk connotes an undetermined residue. The moment a residue is determined, the risk is annihilated." *Op. cit.*, Vol. XV, p. 78.

[3] On this point, see J. R. Hicks, The Theory of Uncertainty and Profit, *Economica*, 1931, pp. 171-189. In this important article Dr. Hicks tries to explain in what different ways entrepreneurs actually may face the fact of uncertainty in the periodical adjustment of their and others' resources.

Clark's conception of economic dynamics is itself worked out against a static background and is much too mechanical. The 'static standard' that rules in the 'dynamic economy' is a sort of a 'norm' which the economic system ever strives to attain. This indeed is a study of what is today known as Comparative Statics. The emergence of profit is considered to be merely a stage in the regular progress of the economy. In the ultimate analysis it turns out to be a 'frictional surplus', the significance of which is just to raise the economic standard. Clark thus misses the active role of uncertainty and anticipation in shaping the course of things. As a matter of fact, a mere conformity of the *external* environment to the static model envisaged by Clark does not ensure the realisation of a correspondence of factor incomes to marginal productivity, or of prices to costs. The objective facts of life may remain stationary,—there may not occur any of those 'five generic changes',—yet, if the economic subjects fail to gauge things correctly, a dynamic element enters into the economic system. Anticipations are in fact a part of those data that shape the economic system. The system cannot be stationary unless the economic subjects *know* that it is to be so.

5. It is, however, a disconcerting paradox that, on the one hand, economic equilibrium is a tendency of a process through time—given the relevant psychological and technical data it takes time for the economic forces to work themselves out to rest,—on the other hand, the longer the time that is held in view, the greater is the possibility that external forces should intervene and obstruct the operation of the original forces. Uncertainty is thus a permanent feature in the economic system. It is one of the limitations of human ingenuity that it cannot unearth the contents of the future. Trained instincts of business men coupled with statistical information may go a long way ; but in so far as

the course of nature (both physical and human) is anything but rhythmical, the future would always remain more or less of a mystery.

Moreover—and this is more significant—the tendency of the economic forces is not necessarily *towards* equilibrium;—the path followed by the different parts of the economic system is not necessarily 'convergent'. Static analysis of the 'functional' type concerns itself mainly with what may be described to be the end-point of economic activities, and pays little attention to the problem of the 'path'. The procedure would be legitimate if the final point were independent of the intermediate points in the adjustment process. But there is no guarantee that it should be so.[1] It is quite possible,—nay, probable,—that the way the forces operate through time should be such as to lead to "a kind of perpetuum mobile of changes". Even apart from exogenous forces, the interaction of the original forces that we start with may be such that they generate over time conditions that lead the system *away from* the so-called static equilibrium. "The swings of the economic pendulum are as likely to grow in amplitude as to diminish. ... The fact that an equilibrium position is attainable, which, if attained, would satisfy all parties, *may* ensure that that position will in fact be attained. On the other hand, there may come about a a perpetual oscillation backwards and forwards round that position."[2]

Now, when positive profits of some entrepreneurs are just balanced by negative profits of others, there emerges no general surplus in the economic system. Net profit as a general category of income arises when such cancellation is

[1] For an exhaustive criticism of the Functional Theory of Prices, see Hans Mayer, "Der Erkenntniswert der functionellen Preistheorien", *Wirtschaftstheorie der Gegenwart*, Bd. 11, pp. 129-239.

[2] Pigou, *Economics of Stationary States*, p. 15. For an analysis of the precise conditions in which the forces are *equilibrating*, see Rosenstein-

not effected, and this happens when there is a relative scarcity of agents ready to bear uncertainty. Under conditions of free competition, as we have seen, the income of the hiring factors tends to be equal to that which they might earn as hired factors ; for, in case any discrepancy arises there is a tendency to a shift of position. It is in equilibrium conditions that the factors receive the full value of their economic contribution to production and there is a profitless distribution of national income. So long as entrepreneurs start operations with imperfect knowledge about the state of the market and the anticipated marginal product of the hired factors deviates from their actual product, so long a surplus would persist.

It is of course difficult to say on *a priori* grounds how the economic subjects would react to the fact of uncertainty. The condition of supply of entreprenours depends upon the extent to which people are willing to take chances. If, however, as is often found, the psychology of the people is such that they are attracted towards high risks,—that a good many of the business men are drawn towards games of chance in which the stake is heavy rather than be content with a steady but moderate income, then, over any fairly long period—*and barring monetary disturbances*,—a positive balance, if any, accruing to the entrepreneurs as a class is not likely to be very large.[1]

Rodan, The Role of Time in Economic Theory, *Economica*, Feb. 1934, pp. 88-97 ; also N. Kaldor, A Clasificatory Note on the Determinateness of Equilibrium, *Review of Economic Studies*, Vol. 1, No. 2.

[1] Cf. Hicks, The Theory of Uncertainty and Profit, *Economica*, May 1931 ; Knight, *op. cit.*, Chap. XII, pp. 362-366.

CHAPTER X

MONEY AND DYNAMIC SURPLUS

1. The complexity of the process is heightened and economic forces tend to produce cumulative effects when changes come through money, and not merely through what may be called 'real factors'.

We have seen that in so far as the demand for one good comes from the supply of another, profits and losses from the point of view of particular industries are relative concepts, so that in a period of change and miscalculation profits in one industry are balanced by losses in another and a general surplus is a misnomer. But of course in actual practice—immediately at any rate—the demand for all goods comes from the holders of money. And in a system in which one particular good (whatever it is) is taken as a 'unit of account', that is to say, as a standard in terms of which the values of all goods are expressed, a general miscalculation leading to an absolute surplus is possible in any economic period if a change occurs in the standard itself. The realisation of the anticipations of entrepreneurs depends not only upon the behaviour of 'real factors' (e.g., the relative valuations put upon goods by economic subjects), it also depends upon the behaviour of money. In fact, money being the unit of account, a change in its quantity may upset the entire structural relation of goods independently of other data.

It should be remembered, however, that a mere change in the quantity of money—if money is only a unit of account—does not bring about any change in the economic relationship of 'real factors'. For, if the change in the quantity of money is foreseen, then all calculations are made with due consideration of the change. What happens is only a

change in the standard of reference ; the deeper relationship of goods remains unaffected. If all of us wake up tomorrow to find that our money incomes have doubled, we shall only pay two rupees for a good for which we had been hitherto paying one rupee. But the fundamental economic situation—the 'relative prices' of goods—would remain unaltered. This is the sense in which the classical economists considered money to be a 'veil'. Taken in this light, more or less of money—high or low 'general price level'—is immaterial so long as the 'real value' hidden under the veil remains the same.

But in fact the behaviour of money is not as simple as that. The conception of money as a veil overlooks the very *raison d'etre* of the institution of money. The main function of money is to bridge up the gulf between the present and the future. If the adjustment of the value relationship between goods and of the distribution of productive resources had been instantaneous,—if there were an omniscient hand manipulating the entire economic scheme and fitting up the different elements into their proper places,—money would be a superfluous institution. Year in and year out streams of goods and services would be pouring into the economic system, just suiting the demands of consumers. There would be no demand for money as such.

But, as we have noted in the last chapter, adjustment of economic forces is a question of time. Each entrepreneur, during any period of planning, has to calculate in terms of a common unit the cost of his undertaking and the anticipated value of it. Money as unit of account is required for this calculation. On the other hand, as soon as the element of time and uncertainty is allowed, the question of demand and supply of money becomes important. It becomes an independent factor affecting the calculation of entrepreneurs. Both the demand and supply of money are regulated

by factors over which entrepreneurs have hardly any control and which they can hardly foresee. The demand is a function of the desire on the part of the public to hold money against unforeseen contingencies. The supply is a function of the eagerness of banks to sell credit. Here is therefore a potent factor causing change in the economic system, the nature of which is different from that considered so far. Money being the unit of account, uncertainty about its behaviour may result in a cumulative tendency away from the position postulated in equilibrium analysis.

This element of uncertainty as invariably associated with money economy has been receiving increasing recognition in recent years.[1] It is coming to be realised that the existence of the use of money and static equilibrium are incompatible phenomena.[2] The problem of monetary

[1] That money is more than a veil was recognised by early writers, too. Cf. particularly, Malthus: "Theoretical writers in Political Economy, from the fear of appearing to attach too much importance to money, have perhaps been too apt to throw it out of consideration in their reasonings. It is an abstract truth that we want commodities, not money. But, in reality, no commodity for which it is possible to sell our goods at once, can be an adequate substitute for a circulating medium, and enable us in the same manner to provide for children, to purchase an estate, or to command labour and provisions a year or two hence. A circulating medium is absolutely necessary to any considerable saving; and even the manufacturer would get on but slowly, if he were obliged to accumulate in kind all the wages of his workmen. ... The circulating medium bears so important a part in the distribution of wealth and the encouragement to industry, that it is hardly ever safe to set it aside in our reasonings, and all attempts at illustration, by supposing advances of a certain quantity of corn and clothing, instead of a certain quantity of money, which every year practically represents a variable quantity of corn, cannot fail to lead us wrong." *Principles* (2nd ed.), p. 324 n.

[2] Cf. P. N. Rosenstein-Roden, The Co-ordination of the General Theory of Money and Price, *Economica*, Aug. 1936, where it is discussed how from two independent approaches—the Cash Balance theory and Capital theory—economists have reached the conclusion about the incompatibility of money economy and static equilibrium. This conclusion which seems at

theory is considered to be "largely one of deducing changes in anticipations from changes in objective data which call them forth."[1] It is with one particular aspect of that theory—the effect of monetary changes upon the profits of entrepreneurs—that we shall be concerned here. We shall see how in an exchange economy such changes may create an absolute monetary surplus in the economic system, and how such surplus may turn out to be a potent engine for influencing the anticipations of entrepreneurs and for causing cumulative disequilibrium in the economic system.

2. It has been noted in the last chapter that the classical economists— Say, Mill and Ricardo, for example, —started with *given* resources (or rather with a condition of full employment), and argued that as goods pay for goods, general over-production is a myth,—that profits accruing to one group of industries are necessarily cancelled by losses in others. This inner relationship is only concealed by money. Money incomes of entrepreneurs are mere external manifestations. Translated in terms of goods a general surplus is a misnomer.

These propositions, however, did not go without challenge. Malthus in particular took the view that there might be such a thing as 'effective demand' in general falling behind supply. "If commodities were only to be compared and exchanged one with another, it would then be true that if they increased in convenient proportions,

first sight to have a somewhat revolutionary appearance should not, however, be understood to have proved the futility of static analysis. The fundamental relation of interdependence of economic forces which is the basic point in static analysis remains. What is suggested is the introduction of some new variables into the scheme,—variables concerning anticipations which may modify the equilibrating tendency over any particular period.

[1] J. R. Hicks, A Suggestion for Simplifying the Theory of Money, *Economica*, Feb. 1935, p. 13.

they might, whatever was their increase, preserve the same relative value. But if we compare them, as we ought, with the number and wants of consumers, a great increase of production with a stationary number of consumers, and wants reduced by parsimony, would of necessity occasion a great fall in the value of the production estimated in labour, so that the same production which costs the same labour as heretofore could no longer purchase the same quantity of it."[1] Malthus thus recommended, as against the then prevailing tendency of extolling the virtue of saving, an increase of 'unproductive consumption' just for keeping up 'effective demand'.

Now, Malthus' analysis of 'effective demand' is no doubt crude. It implies a comparison between wants of consumers taken in an absolute sense and absolute labour cost of goods, to which no precise meaning can be attached. Secondly, he fails to adduce sufficient reasons why consumption should lag behind production. Indeed one might feel amused to find the author of the *Essay on Population* scenting the danger of stationary population![2] Yet this controversy between Malthus and his opponents— Say and Ricardo—involve issues which are of the highest importance in economic theory and practice. The arguments of Malthus contain elements which are essential for dynamic analysis in money economy. The Say-Ricardo proposition is appropriate for the more abstract stationary equilibrium analysis where money is not only a 'veil' but is strictly speaking

[1] *Principles of Political Economy* (1820), p. 355; 2nd edition (1836), p. 317, where the language is slightly altered.

[2] This paradox was pointed out by J. B. Say in one of his letters to Malthus. In connection with this interesting controversy on what is known as the Theory of Gluts, read, in addition to the books already mentioned, J. B. Say's *Letters to Malthus*, esp. Nos. 1-3, published in *The Pamphleteer*, No. XXXIV; also Ricardo, *Notes on Malthus*. A good discussion on this controversy is to be found in Bonar, *Malthus and his Works*, Bk. 11, Chap. III; See also Keynes, *Essays in Biography*, pp. 137-145.

a non-entity.[1] In money economy an absolute surplus in the shape of profits and losses may arise in the economic system, the potency of which was ignored by Say and Ricardo. In fact at the criticism of Malthus, Say had ultimately to restrict the application of his theory and base it on an assumption which, as Dr. Rosenstein-Rodan has pointed out,[2] may be said to have been an anticipation of what is today known as a condition of 'monetary equilibrium',—a condition, that is to say, of equality between saving and investment in the Wicksellian sense.[3]

In his theory of *Interest and Prices* Wicksell, it will be remembered, raises the question of how the effective demand and hence the prices of *all* (consumers') goods may have a tendency to rise or fall, and how taking *all* goods together there may arise profits and losses. Wicksell's saving-investment theory and the theory of monetary equilibrium proceed from a realisation of the importance of money as more than a veil. "Economists frequently go too far when they assume that the economic laws which they have deduced from barter assumptions may be applied without qualification to actual conditions, in which money effects practically all exchanges and investments or transfers of capital.... The

[1] As Mr. Keynes points out, "Ricardo is investigating the theory of the *distribution* of the product in conditions of equilibrium, and Malthus is concerned with what determines the volume of output day by day in the real world. Malthus is dealing with the monetary economy in which we happen to live; Ricardo with the abstraction of a neutral money economy". *Op. cit.*, p. 138.

[2] *Op. cit.*, pp. 268-269.

[3] In one of his letters to Malthus, Say admits the possibility of saving being unproductive. "As to sums accumulated without being productively consumed, for instance, those hoarded up in the miser's coffer, neither Smith, myself, nor anyone undertakes to defend this, but they alarm us but little in the first place, because they are very inconsiderable in comparison to the productive capitals of a Nation, and in the second place because their consumption is no more than suspended." *Op. cit.*, p. 316.

use—or the misuse—of money may, in fact, very actively influence actual exchange and capital transactions."[1] He thus gives a description of 'monetary' equilibrium (as distinct from the stationary equilibrium of 'relative prices'),—a tendency away from which leads to absolute profits and losses, which again produce a 'cumulative process' in the economic system. In barter analysis, frictions causing profits and losses are shown to be temporary, and the economic system is shown to be moving round a centre. Monetary changes not merely alter the 'form' of the movement but also its 'essence'. The possibility of positive or negative profits "arises out of the fact that the transfer of capital and the remuneration of factors of production do not take place in kind, but are effected in an entirely indirect manner as a result of the intervention of money. It is not, as is so often supposed, merely the *form* of the matter that is thus altered, but its very essence."[2] Wicksell compares the static equilibrium of relative prices with a pendulum which, if disturbed, would oscillate round a point and would come to rest again. But the movement and equilibrium of money prices exhibit a fundamentally different character. It is compared with a cylinder resting on a rough horizontal plane 'in a so-called neutral equilibrium'. The impact of a force sets the system of money prices moving in the same direction : "the motion is an accelerated one up to a certain point, and it continues for a time even when the force has ceased to operate. Once the cylinder has come to rest, there is no tendency for it to be restored to its original position. It simply remains where it is so long as no opposite forces come into operation to push it back."[3] The element that brings this whole 'cumulative process' into operation is

[1] *Lectures on Political Economy* (ed. Robbins), Vol. II, p. 6.

[2] *Interest and Prices*, p. 135.

[3] *Ibid*, p. 101 ; cf. also *Lectures*, Vol. II, pp. 196-197.

profit, positive or negative. The condition of monetary equilibrium, or in other words the absence of cumulative process, is therefore the condition of zero profit.

Wicksell gives three conditions of this equilibrium: (i) equality between the loan rate and the 'natural rate' of interest; (ii) equality between the supply and demand of real capital; and (iii) stability in the price level. The third condition has been disposed of by later writers as being not necessarily consistent with the other two. It is now perhaps universally accepted that even below a stable price level there may be developing serious tendencies to mal-adjustment when the economy under consideration is a growing economy.[1] The first two conditions of monetary equilibrium, or of zero profit, resolve themselves to this, that if the banking policy is regulated in such a way that the actual money rate of interest corresponds to the current value of the marginal productivity of real capital, then the economic system keeps in equilibrium, or in other words, even under money economy the system moves as if it were a barter economy, and the formation of capital proceeds just as if it were a transference of resources *in kind* from lower stages to higher stages of production. If, however, these conditions are not satisfied, profits or losses arise and the

[1] On this point, see Hayek, *Prices and Production*, *passim*. For an analysis of the American experience confirming the view that a stable price level may be incompatible with a stable economic system, see Robbins, *The Great Depression*, Chap. III. 7. This criticism of Wicksell's third condition of monetary equilibrium was first made by Wicksell's own contemporary, Davidson. See Brinley Thomas, The Monetary Doctrines of Professor Davidson, *Economic Journal*, March 1935. "When productivity increases, the requirement that the value of money should remain stable demands that prices should fall: When the production of goods increases without a rise in productivity, the same requirement demands that the price-level should remain stable." (Davidson). Quoted by Dr. Thomas; *op. cit.* p. 38.

cumulative process sets in. If the money rate of interest is kept below the 'natural rate', there is positive profit which induces entrepreneurs to invest more. "The producer has to pay more for raw materials, wages, rents, etc., but he receives correspondingly better prices for his own products. He finds himself in precisely the same situation as before the rise in prices took place, and he is therefore in a position to pay the same rate of interest as before for the credit which he requires. If, however, the credit institutions maintain the lower rates of interest, he will be in a position to offer rather more for raw materials, labour, and land, and competition will to some extent force him to do so. As a consequence, the demands of workers and landlords will be raised, and they will bring about a further rise in the price of consumption goods ; and prices will continually rise higher and higher."[1] The process continues until ultimately banks are compelled to raise the market rate of interest to the level of the natural rate. A reverse process operates when the market rate is kept above the natural rate.

3. This requires further elucidation. In a simple economy, where the saver is himself the investor and where there is no intervention of money, an individual starting with given resources in terms of time and energy has necessarily to withdraw from consumption an amount which he decides to put in as new investment. What is saved consists of real resources, and an act of new investment is merely a transference of real resources from immediate consumption to future consumption. If the decision is made with full realisation of his subjective preference for present and future wants, then this transference puts the individual in a new equilibrium position. If, on the other hand, the decision is made with inadequate foresight, he may find at the end of the operation that the value of his present wants

[1] *Interest and Prices*, p 95.

is unduly raised *in relation* to the value of his future wants, so that a revision is called for.[1]

In money economy, however, things are different. Here the process of capital formation is not so direct. The intervention of money makes the thing complex. Savings and investments are made by differents groups of people. Transference of resources from consumption to capital formation is effected through the banks offering credit to entrepreneurs, and the regulating factor is the rate of interest. Now, according to Wicksell, there is a rate of interest at which just that amount of resources will be diverted to the production of capital goods which is warranted by the volume of saving of the community. This rate of interest, "the natural rate", is equal to the current value of the (physical) marginal productivity of capital. If the market rate coincides with this, the entrepreneurs are in a position to draw just that amount of resources which is released on account of a withdrawal of consumption in the form of savings of the community. This is the condition of monetary equilibrium. It is a state of things in which harmony is kept up between the proportion in which resources are employed by entrepreneurs as between consumption goods and capital goods and the proportion in which individuals choose to distribute their income between consumption and saving. It is a state, in other words, in which, although savings and investments are made by different sets of persons, banks act as an intermediary and bring the two in equilibrium. The system works as if resources were directly transferred through voluntary abstinence of the community to the formation of capital. If, however, the market rate is kept below the natural rate, entrepreneurs as a whole

[1] It should be noted that the factor of uncertainty is there even in the so-called Crusoe economy. Even an isolated man may not be able to foresee the effect of a particular action upon his own susceptibilities.

enjoy positive profits and are induced to invest more; and if the market rate is kept above the natural rate, entrepreneurs as a whole suffer loss and cut down investment. These processes of change—the movement of the economic system *as a whole* forward or backward—continue so long as there is any divergence between the two rates.

Now, although, as we shall see, this analysis of Wicksell is open to a serious objection, it provides a clue to the study of economic dynamics. It provides elements round which modern analyses of the processes of change in the economic machinery have centred.

4. Wicksell defines 'natural rate' of interest as the current value of the *physical* marginal productivity of capital,—a rate of interest which would evolve if there were no money and no credit. But physical marginal productivity is a vague expression when we are concerned with heterogeneous kinds of capital goods and heterogeneous products. In order to have an idea of the physical marginal productivity of capital, not only have we to assume that resources are homogeneous, we have to assume further that the yield is homogeneous with the source. Wicksell knew it.[1] But he tried to overcome the difficulty by assuming an invariable exchange relation between factors and products. This, however, is assuming away the problem. For the exchange relations that are taken to be given and invariable are themselves dependent on 'time preference'. In fact, as Myrdal points out, "credit and interest rate must be included in the very construction through which the natural rate is determined."[2] Barter

[1] ".. Capital, like loaned money, has the peculiarity that its share in the product—interest—is the *same kind* of thing as capital itself; interest is an organic growth out of capital, a certain *percentage* of capital." *Lectures*, Vol. 1, p. 145.

[2] Der Gleichgewichtsbegriff als Instrument der geldtheoretischen Analyse, *Beitrage zur Geldtheorie* (ed. Hayek), seg. 27-28. For a criticism of the concept of natural rate, see also Sraffa, Dr. Hayek on Money and Capital, *Economic Journal*, March 1932.

analysis and natural rate in the Wicksellian sense belong to a realm of unreality. They may be useful as a starting point, just as Adam Smith's parable of beaver and deer is useful as a starting point. But in an exchange economy, where we have to deal with heterogeneous factors and heterogeneous products, 'natural' rate can only be conceived in terms of money.

As, moreover, the yield of investment is a flow through time, at any moment of planning, the 'natural' rate to be taken as a standard for comparison in considering the rate of investment can only be interpreted as the *expected* rate of yield. Myrdal designates it as *Extragsquote des Real-kapitals*, which corresponds to Mr. Keynes' Marginal Efficiency of Capital,—a rate of interest which would equate the present value of the expected net yield of a capital asset in future periods with its supply price, both in money terms.[1]

Thus in calculating the value of capital,—in other words, in discovering the capitalised value of the future net receipts from investment,—entrepreneurs have to take the actual rate of interest into account. Natural rate thus defined, therefore, loses all the sharpness that it had in Wicksell's system. For a conformity of the actual rate of interest with the marginal efficiency of capital—towards which surely there is always a tendency—does not imply a condition of 'monetary equilibrium' in any normative sense. It is just a stage in the movement of the economy.

It is expected productivity of capital at the margin together with the current rate of interest that decides the volume of investment, and it is the volume of net profit emerging in the process that is the mainspring of change in the economic system. If the expected net yield of investment exceeds its cost, the entrepreneur expects positive profits and is induced to invest more. If, on the other hand,

[1] Cf. *The General Theory of Employment, Interest and Money*, p. 135.

the cost of investment exceeds its value, there is depression in investment activities. Now, given the market rate of interest, the entrepreneur's expected profit depends upon (*i*) his expectations concerning gross incomes during the lifetime of the capital good, (*ii*) expectations concerning gross costs during the same period and (*iii*) expectations concerning changes in the capital value itself, this latter involving expectations concerning the future course of interest rates,—all these again being corrected with reference to the probable degree of risk.[1] These are what the Swedish economists call *ex ante* calculations. If these expectations prove correct at the end of each period, profit *ex ante* coincides with profit *ex post*. But they may not be—indeed are not likely to be—correct. If they are not, then from each of these sources positive or negative profits may result. Present expectations of profit influence the present conduct of entrepreneurs; present conduct determines the profit to be realised in future; these realised profits again form *one of the variables* that determine future expectations and future conduct. Economic process is thus a process of "shifting" or "dynamic" equilibria, each point of equilibrium being characterised by an equality between *value* and *cost* of investment.

Now, in what way actually does a dynamic surplus in the form of net profit arise in the general system, and what precisely is its behaviour?

An initial investment activity, whatever it might be due to, may be financed by the savings of the community, or by banks creating new credit. The total circulation of money remains constant in the former case, and increases in the latter. In the former case, all that the process amounts to is a redistribution of resources based on the relative profitability of capital-goods industries and consumption-goods

[1] These calculations thus vary with the kind of investment and with the judgement of individual entrepreneurs.

industries. The income of the community remains the same, and the static hypothesis holds good. When, however, banks lower the rate of interest and expand credit, entrepreneurs as a whole secure an unexpected profit. For, if the expansion means higher factor prices, it also means higher commodity prices and higher income. If, therefore, the entrepreneurs expect this state of affairs to continue, Wicksell's cumulative process sets in, and the entire economic system starts moving forward. Dynamic disequilibrium is thus characterised by a divergence between the expected and the realised profit of entrepreneurs as a a whole.

The behaviour of this dynamic surplus, however, is not unique and determinate through time. The profit-earners are supposed to be the principal savers of the community. The creation of this abnormal profit therefore leads to a growth of savings at an increasing rate. Do these savings find their way to investment in real capital and push the economic system forward,—or do they, by bringing about a relative reduction in the price of consumption goods, impair the expectation of entrepreneurs and depress investment activity? In the former case, a low rate of interest is a source of an over-expansion of investment, whereas in the latter case, the continued maintenance of a low rate of interest is a condition necessary for keeping up investment and employment.

It is here that the role of anticipation comes in. For, as we have seen, the volume of investment is influenced by the expectation on the part of entrepreneurs about the marginal productivity of capital. This expectation depends, among other things, upon the particular stage of capital accumulation and economic development. It is probable that during an early stage when there prevails a wave of optimism, a low rate of interest which gives rise to profit would create further expectation of profit and would provide

a fillip to investment. But after a time when there is a large accumulation of capital, a reaction would set in, expectation would be damped and investment discouraged except at still lower rates of interest.

5. These considerations have played an important part in modern discussions on Trade Cycles. The capitalist economy is found to be characterised by ups and downs in business activities. These 'booms' and 'depressions' in trade and industry are not mere accidental aberrations, but appear to be in the nature of things. Clearly, they are the result of wrong anticipations on the part of entrepreneurs. If producers and consumers had perfect foresight, the economic system would have a steady course ; the production plan would just adjust itself to the consumption plan.[1] With uncertainty present, disturbances are bound to occur. And in seeking for an explanation of a general disturbance one should naturally turn to the behaviour of money.

Wicksell himself did not apply his saving-investment theory to the explanation of Trade Cycles,[2] although in a Note on Trade Cycles he recognised that monetary disequilibrium does accentuate industrial fluctuations. It is suggested there that "changes in the purchasing power of money caused by credit are under existing conditions certainly bound up with industrial fluctuations and undoubtedly affect them, especially in causing crises, though we need not assume any necessary connection between the phenomena."[3]

[1] Planning of consumption has its dynamic aspect, too, no less significant than the dynamics of production plan. Uncertainty in consumers' expectations about the future, which keeps them indefinite about the future way of spending resources and which therefore influences their 'liquidity preference', is an important factor causing disturbance. See Rosenstein-Rodan, The Role of Time in Economic Theory, *Economica*, Feb. 1934, Sec. 3, pp. 80-84 ; also Hicks, *Value and Capital*, Chap. XVIII.

[2] See Ohlin's Introduction to *Interest and Prices*.

[3] *Lectures*, Vol. II, p. 211.

Two outstanding developments in Trade Cycle theory deriving from Wicksell's propositions are those of Professor Hayek and Mr. Keynes.[1] And recent discussions on Trade Cycles have centred on their theories. Now, up to a point Keynes' analysis of the problem—at any rate in the *Treatise*—runs parallel with Prof. Hayek's. Both derive from the theoretical foundation laid by Wicksell. Both diagnose disequilibrium to be a mal-adjustment between the market rate of interest and the 'natural rate',[2] or, to be more precise, as owing to the non-neutrality of money. 'Windfall profits' of the *Treatise* can in some sense be taken to correspond to 'forced savings' of Prof. Hayek. But the emphasis put upon the different elements in the system by the two authors is different.

Hayek points to the danger of an excess of investment over what he calls 'voluntary' savings as creating an un-

[1] These two names are particularly mentioned because Prof. Hayek represents the more rigorous type of what is called the Monetary Over-investment theory of Trade Cycle, while its opposite, the Under-consumption theory, finds its culmination in Mr. Keynes' analysis. For an exhaustive analysis of the different theories of Trade Cycle, see G. Haberler, *Prosperity and Depression* (1937), Part I.

[2] It should be noted that even in the *Treatise* Mr. Keynes, unlike Wicksell, conceives the natural rate to be essentially a money rate, and he also argues that a disturbance in monetary equilibrium is corrected not necessarily through an adaptation of the market rate to a rigid natural rate, but that a raising or lowering of the market rate may set up forces which move the natural rate itself towards the market rate. A rise in the rate of interest, which means a fall in the price of new capital goods "must necessarily be deterrent to the production of such goods, until, as a result of it, the falling off in their prospective supply has raised the money-value of their prospective yield sufficiently to offset the effect of the higher rate of interest". (*Treatise*, Vol. 1, p. 203). The natural development of this is that monetary equilibrium is in essence what Dr. Rosenstein-Rodan calls 'dynamic equilibrium'. (Cf. Co-ordination of the Theory of Money and Prices, *Economica*, August, 1936.). Indeed, in formulating the concept of natural rate, Mr. Keynes may be said to have been feeling his way towards what he now calls the 'marginal efficiency of capital'.

healthy boom, which, once started, inevitably leads to a crisis. An expansion of credit on the part of banks through a rate of interest lower than the 'equilibrium' rate opens out prospects to entrepreneurs dealing in capital goods,—those goods, that is to say, that take a longer time for production and as such are more affected by the price of 'waiting'; but the prospects opened out are more ephemeral than real and progressively lead the activities of entrepreneurs away from equilibrium. If it persists, it produces such cumulative errors in the expectation of entrepreneurs and gives such a turn to the economic system that ultimate adaptation becomes a long and painful process. An artificially low rate of interest raises the profitability of capital goods and tends to draw resources from lower stages to higher stages of production. The structure of production is 'elongated'. But a time soon comes when there arises a shortage of liquid capital ; and the top-heavy structure thus created cannot be maintained. When the effect of the expansion of credit is transmitted into the system, the rate of interest rises and the structure of production has to be contracted.[1] In the process of this contraction there are business failures. The enterprises launched with high prospects fail to be kept up. The error of optimism asserts itself into a collapse and a depression.

Mr. Keynes would not lay so much emphasis on the distinction between 'voluntary' saving and 'forced' saving. In the *Treatise*, as we have noted, 'windfall profit' has some character of forced saving, although Keynes himself would perhaps deny this analogy. In the *General Theory* he has explicitly repudiated it. Mr. Keynes, on the other hand, is afraid of an excess of saving which by itself might bring in

[1] If the credit offered is producers' credit, the system goes through this process. If it is consumers' credit, the contraction of the structure of production takes place immediately.

the depression. According to him an excess of saving brings down the prices of consumption goods and hence the capital value of resources, thus toning down the inducement to invest.[1]

Prof. Hayek's analysis assumes that the demand for resources from consumption-goods industries on the one hand, and capital-goods industries on the other, is a *competitive* demand. Although from the point of view of a single industry the demand for machines is *derived* from the demand for their products,—from the point of view of the economy as a whole the relation is not complementary. Mr. Keynes suggests that if there is any relation between them, it is that they more often than not tend to move in the same direction. According to Prof. Hayek there is a mechanism—the rate of interest—through which the saving of the community finds its way to investment. According to Mr. Keynes there is no such mechanism, for it all depends upon the individuals' scale of preference between spending and holding cash.

All these divergences are to very large extent traceable to a difference in the fundamental assumption upon which they proceed. Prof. Hayek starts with a position of static equilibrium in which there are no unused resources and explains trade cycles as a tendency away from such equilibrium.[2] Mr. Keynes, on the other hand, assumes a condition

[1] The mode of expression is that used in the *Treatise*. In the *General Theory* the 'excess saving factor' and the 'excess bearish factor' of the *Treatise* boil down to increased 'liquidity preference' and fall of income.

[2] "If we are to proceed systematically .. we must start with a situation which is already sufficiently explained by the general body of economic theory. And the only situation which satisfies this criterion is the situation in which all available resources are employed. The existence of unused resources must be one of the main objects of our explanation." *Prices and Production*, p. 35. See also *Monetary Theory and the Trade Cycle*, Chaps. I-II.

of under-employment and then asks what follows if there is a change in investment activities.

Now, it is true that if stationary equilibrium is incompatible with money economy, as indeed Prof. Hayek's own analysis implies, it is not legitimate to start with an equilibrium position. And once we start with unused resources, some of Prof. Hayek's conclusions do require modification. Moreover, this original assumption of full employment is responsible for the fact that he did not analyse as closely as he otherwise would have done the important role of uncertainty in the dynamic process. In the third place—and this also follows from the problem of uncertainty—he did not pay sufficient attention to the aspect of liquidity preference. Taking an initial position of disequilibrium it is easily seen that for a time a healthy progress can be maintained through an expansion of credit, which would lead to what may be called an induced saving.

Yet it cannot be denied that at a time when it was going to be almost a fashion to take a stable price level as the panacea for all economic disturbances, Prof. Hayek was among the few who put their finger on an important disequilibrating element in the economic system,—the vagaries of the rate of interest. It is now a commonplace that the real causes of a crisis are to be traced to a preceding boom, and that the intensity and duration of the depression are largely a function of the extent to which the boom had been carried. It is also generally recognised that booms and depressions are characterised more by fluctuations in the activities in capital-goods industries. A period of increased investment leading to an expansion of capital-goods industries is followed by a contraction; and the process from boom to depression is marked by business failures which appear in the form of a crisis. Now, psychological forces do play a most important part in the cycle of

booms and depressions. It is not facts but opinions about facts that guide human conduct. And opinions have a momentum-gathering virtue. But opinions do not run *in vacuo*. They are reactions to objective facts of a situation. A general turn of anticipation concerning future prices must be sought, in part at any rate, in the actual price-cost-relationship obtaining at the moment; and any sustained mal-adjustment between prices and costs leading to a cumulative error of anticipations is to be explained partly by changes occurring on the side of money. Whatever may be the initial cause of a general optimism in favour of investment, the investment boom is not likely to persist unless it is backed by an expansion of credit. It is thus difficult to understand Mr. Keynes' indifference to the part played by a low rate of interest in the boom period,— or his denial of the scarcity of capital.[1] If resources were indefinitely elastic, then of course the problem would be simply that of keeping up the spirit of enterprise,— of raising the 'marginal efficiency of capital'. But if they are not elastic, then certainly a low rate of interest in times of boom would only tend to *divert* resources to channels which would ultimately turn out to be unprofitable.[2]

[1] Cf. *General Theory*, p. 376. "Interest today rewards no genuine sacrifice, any more than does the rent of land. The owner of capital can obtain interest because capital is scarce. But whilst there may be intrinsic reasons for the scarcity of land, there are no intrinsic reasons for the scarcity of capital."

[2] Cf. Wicksell, *Interest and Prices*, p. 143. "It is impossible to endorse the widespread view that under suitable conditions a country's output can be expanded almost indefinitely by 'arousing the spirit of enterprise' and the like. This fallacious view is derived by concentrating attention on one single branch of production, provided perhaps with an excess of fixed capital (buildings, machinery, etc.). In such a single branch of production it would be possible to increase output immediately, but only at *the expense of the other branches of production* from which labour and liquid capital have to be drawn."

But while it is true that an over-excitement of entrepreneurs backed by a low rate of interest may cause a crisis, the after-math of such crisis, the 'secondary depression', as it is often called,[1] is characterised by an increasing propensity to 'hoard' and a collapse in the expectation of entrepreneurs in regard to the prospective yield of investment. If an investment boom leads to a crisis, the crisis is accentuated and recovery of the system becomes all the more difficult in view of a growing uncertainty that characterises the situation,—a growing uncertainty which raises the demand for money on the part of the public, and thus keeps up the rate of interest; and this at a time when the expectations of entrepreneurs are at a low ebb. This leads to cumulative unemployment, which again tends to persist so long as business confidence is not revived. It is here that Mr. Keynes' analysis is most illuminating.

6. Now the question arises, if uncertainty is a permanent feature in our economic life, and if, as we have seen, it necessarily leads to fluctuations in trade and industry with accompanying profits and losses, is it to be left to be borne by private entrepreneurs or should it be undertaken by the State? Happenings of the recent past—the great cycle of boom and depression—have given rise in certain quarters to a disbelief in the efficiency of the individualist exchange economy based on private enterprise and cries have been raised against what is described to be an 'anarchy' in production. Very likely this tendency will be gathering more strength during the years to come.

It is true that major fluctuations like the one that we experienced after the last war are real evils to be eradicated, and economists have been giving their closest attention to the diagnosis and probable remedies of them. It may be hoped that with a little more judicious use of money, and a

[1] Cf. Röpke, *Crises and Cycles*, pp. 119-133.

wider dissemination of knowledge much progress will be made in that direction in future. One ought to remember that the choice is not simply between 'planning' and a crude and ill-conceived *laissez faire* policy. The universe of discourse is not exhausted by just these two alternatives. There is the third alternative which raises the problem of the degree and direction of interference. Modern analyses in economic dynamics have revealed possibilities of a judicious interference of the State and the monetary authority towards controlling the wayward expectations of speculative investors. Statistical studies in this direction may help matters more. But fluctuations there must be so long as the future remains at all a blurred leaf. And it is a debatable question whether the absence of uncertainty would be a desirable condition. Stationary economy is a theoretical construction which merely brings into relief the interworking of forces in the economic system. It is not a 'norm' to be followed in practice.

Uncertainty is not a phenomenon peculiar to the individualist exchange economy. It is inherent in the very nature of economic life. The system of private enterprise is only a method—perhaps the most efficient method yet devised—of grappling with this fact of uncertainty. Even if a central planning authority—whatever form it might take—assumes the task of organising production, the same difficulties will be there,—the difficulties of gauging the changing structure of demand and of varying the combination of factors according to their relative scarcity. But the spur of profit and—what is more important—a guide to the calculation of profit will be absent. In fact we misjudge the role of entrepreneurship in modern business organisation if we fail to realise that they are but functionaries of the society through whom the adaptation of the productive mechanism to changing economic

conditions is effected. And it is also a virtue of the competitive system that the entrepreneurs have a tenure of office according to 'good behaviour'. If a principle of substitution with regard to hired factors works directly through the entrepreneurs, a principle of substitution with regard to entrepreneurs themselves works through the society which consists ultimately of consumers. Profits and losses work as an automatic lever through which this principle operates.

Those who want to eliminate this automatic mechanism fail to realise the immensity of the task that in the alternative would fall upon the controlling authority. In fact if fluctuations due to uncertainty are the 'evil' which 'planning' seeks to eradicate, it is also in view of economic change and uncertainty about this change that the difficulty of planning reveals itself most conspicuously.[1] Wisdom demands a more cautious attitude on the matter.

[1] On all this see Mises, *Socialism*, esp. Part II, Sec. 1, Chaps. VI and VII.

INDEX OF AUTHORS CITED

Allen, D., 127n

Bailey, S., 60
Barone, E., 125n
Böhm-Bawerk, E. v., 36n, 75
Bonar, J., 193n
Bowley, A., 131-32n

Cairnes, J., 53, 72-73, 182
Cannan, E., 54, 123
Cantillon, R., 18-20, 181-82n
Cassel, G., 42n, 70, 80
Clapham, J. H., 5n
Clark, J. B., 65, 80, 104n, 182-83, 186

Dalton, H., 105n
Davidson, D., 196n
Dobb, M., 13n, 14n, 71n
Duncan, A. J., x
Dupuit, 9, 108-12, 114-16, 137
Durbin, E. M., 47n

Edgeworth, F. Y., 94, 125n, 128n, 137, 151n
Einaudi, L., 23n, 108n

Fisher, I., 131n, 132n

George, H., 23, 41
Gobbi, U., 120, 123
Gossen, H., 108
Graham, F., x, 149n, 157, 165

Haberler, G., 151n, 204n
Hawley, F. B., 184-85
Hayek, F. v., x, 196n, 204-07
Hicks, J. R., ix, 127n, 128n, 185n, 188n, 192, 203n
Higgs, H., 23n
Hirst, F. W., 48n
Hobson, J. A., 45-47, 121
Hotelling, H., ix
Hume, D., 48

Jannaconne, 102n 104n
Jevons, W. S., 19, 37, 109n, 112-16, 140-42, 153n
Johnson, A. S., 10-11

Kaldor, N., 188n
Keynes, J. M., 193n, 194n, 200, 204-09
Knight, F. H., 43, 72n, 76n, 80n, 123n, 179n, 184, 185n, 188n

Lerner, A. P., 146n
Lloyd, 108

Malthus, T., 4, 26, 59n, 191n, 192-94
Marshall, A., x, 9, 12, 73, 79-82, 104n, 107n, 117-37, 143-53, 156-67, 180n, 182
Marx, K., 6-7, 30-40
Mayer, H., 187n
Menger, C., 108
Mill, James, 179n, 192
Mill, J. S., 40, 42, 59, 107n, 139-40, 153n, 179n, 182
Miller, H. E., 123n, 174n
Mises, L. v., 211n
Myrdal, G., 199-200

Nicholson, 125n

Pantaleoni, M., 102
Patten, S. N., 106-07n
Petty, W., 20
Pigou, A. C., 12n, 187

Quesnay, E., 17, 19

Ricardo, D., 8, 28-29, 53-67, 108, 192, 193-94
Robbins, L., x, 47n, 75n, 77n, 80n, 94, 105, 196n
Robertson, D. H., x, 82n
Robinson, J., 39n, 72n
Röpke, W., 209n
Rosenstein-Rodan, P. N., x, 77n 130n, 176n, 188n, 191n, 194, 203n, 204n

Say, J. B., 21, 179n, 192, 193n, 194
Schumpeter, J., 183-84, 185n
Senior, N., 28n, 53, 69-72, 182
Sidgwick, H., 41
Smith, A., 4, 24-28, 48-53, 107
Sraffa, P., 199n

Thomas, B., 196n
Turgot, 22, 40, 45, 177-78

Viner, J., x, 96, 157, 159-67

Walras, L., 70, 116n
West, E., 28

Whitaker, A., 53, 57
Wicksell, K., 65, 194-97, 199-200, 203, 208n
Wicksteed, P. H., 34n, 65, 72n, 76n, 103
Wieser, F. v., 5, 42n, 49, 76n
Young, A., 36, 148n, 151n, 152n

ERRATA

P. 40, footnote 1, *for* XCVI, *read* XV.

P. 45, line 2, *for* 'arise', *read* 'rise'; footnote 1, *for* XCVI, *read* XCVIII.

P. [illegible], line 11, *for* 'abourers', *read* 'labourers'.

P. 147, line 18, *for* 'form', *read* 'from'.

P. 163, line 6, *for* MT', *read* M'T'.

P. 171, line 11, *for* 'them' *read* 'it'.

P. 201, line 12, *for* 'profit *ex ante* coincides with profit *ex post*', *read* 'income *ex ante* coincides with income *ex post*'; line 15, *for* 'profits', *read* 'surplus'.

Besides these there are certain misprints which, it is hoped, will be too obvious to the reader to need mention.